Anonymous Referee #4

General remarks:

The present manuscript provides a monthly climatology (from 2007 to 2015) of African dust based on an optimised CALIPSO dust product was recently developed with a regional correction of the Saharan dust LR using EARLINET measurements (Amiridis et al., 2013). The monthly climatology of African dust obtained allows the description of the spatiotemporal features of dust properties over North Africa and Europe. The study of the mean state climatology shows strong seasonal shifts in dust source regions and transportation pathways. While the results of the study are interesting to be published, their presentation and discussion are not yet sufficient to be published in Atmospheric Chemistry and Physics in the current form. Therefore, it is worth to be published after addressing major revisions which are explained below along with a few other details.

[REPLY] We thank the reviewer for the thorough revision and comments. Replies to the general and specific comments follow below.

Major comments:

In Amiridis et al. (2013), this EARLINET-optimized CALIPSO dust optical depth (for the period 2007-2011) is described and qualitatively compared with MODIS and AERONET. The present manuscript is focusing on the analysis of the resulting EARLINET-optimized CALIPSO dust climatology. I would be desirable to include a short discussion of the uncertainties of the EARLINET-optimized CALIPSO dust product. I understand that this discussion is partly in Amiridis et al. (2013, 2015) although the authors should include a summary in Sect. 2.2 as well as about the uncertainties of the algorithm of CALIOP to determine the corresponding aerosol subtype (in Sect 2.1).

[REPLY] We added a new section 3.5 discussing the uncertainties related to the EARLINET-optimized CALIPSO dust product. Moreover, we added a summary in Sect. 2.2 about the uncertainties of the algorithm of CALIOP to determine the corresponding aerosol subtype, referring to the evaluation done with NASA's HSRL and presented in Burton et al. (2013).

The summary, (page 4, line 24): "Burton et al. (2013) showed an 80% successful detection of dust from CALIPSO, upon comparison to underflights with the HSRL system of NASA. This score is considered very high for aerosol typing purposes and is attributed to the depolarization measurement capability of the CALIOP sensor."

The new section 3.5 added in the paper (page 6):

"2.4 Dust product uncertainties

The sources of uncertainties for the pure-dust product are discussed in this section. CALIOP is able to detect aerosol layers with $AOD > 0.005$ and $\beta > 0.25\ Mm^{-1}\ sr^{-1}$ (Winker et al. 2009). The uncertainty estimation of particulate backscatter, extinction and AOD retrievals reported in the CALIPSO Level 2, Version 3 Data Release, are based on the simplified assumption that all the uncertainties are random, uncorrelated and produced no biases

(Young, 2010). More specifically, ignoring multiple scattering, the errors in the layer optical depth calculations typically arise from three main sources: (a) signal-to-noise ratio within a layer, (b) calibration accuracy, and (c) the accuracy of the lidar ratio used for the extinction retrieval. The lidar ratio uncertainty is the dominant contributor to the total uncertainties, and the relative error in the layer optical depth is always at least as large as the relative error in the lidar ratio of the layer, and grows as the solution propagates through the layer (CALIPSO L2-V3, 2010). In our dataset the typical uncertainties in the CALIPSO Level 2 version 3 product are between 30% and 100% for the AOD, between 30% and 160% for the aerosol backscatter and extinction coefficient and >100% for the particle depolarization ratio.

Several studies report that CALIPSO underestimates the columnar AOD due to undetected aerosol in the free atmosphere. For instance, Rogers et al. (2014) report a ~ 0.02 AOD CALIPSO underestimation, when compared to collocated airborne HSRL measurements over the North American and Caribbean regions at night. In their data, the dust layers were primarily non-opaque with extinction less than $1\ km^{-1}$ so there were negligible multiple scattering effects. The aforementioned detection limits and uncertainties of CALIPSO products are propagated to the dust product presented here.

As already described, the EARLINET-optimized CALIPSO dust product is derived using the depolarization-based separation method, coupled with the selection of a uniform climatological LR value. These steps introduce uncertainties in the pure dust product. In particular, the uncertainty in the selection of the representative LR (55 ± 11) is 20% for the study area (e.g. Wandinger et al. 2010; Baars et al. 2016 and references within). This uncertainty in LR is less than half of the uncertainty of the generic LR in CALIPSO version 3 product (40 ± 20 for dust layers and 55 ± 22 for polluted dust layers). As already addressed in several studies (e.g. Wandinger et al. 2010; Schuster et al. 2012; Amiridis et al. 2013), CALIPSO V3 dust extinction coefficient and AOD values are about 30% lower than those obtained from collocated ground-based Raman lidar retrievals due to the low LR used in the CALIPSO aerosol retrievals. Amiridis et al. (2013) applied the EARLINET LR for the pure dust CALIPSO cases above North Africa and Europe, and compared with synchronous and collocated AERONET measurements. The results showed an absolute bias on the AOD of the order of −0.03, improving on the statistically significant biases of the order of −0.10 reported in the literature for the original CALIPSO product. The bias of -0.03 is similar to the low bias of CALIPSO's column AOD due to undetected aerosol layers. In Kim et al. (2017), they found a global mean undetected layer AOD of 0.0031 ± 0.052 by comparing 2 year of CALIPSO (L1-V4) and MODIS AODs.

Regarding the error induced from the application of the dust separation method, this might be due to the selection of the particle depolarization ratio of dust and the other aerosol types (marine, anthropogenic or smoke). Tesche et al. (2009; 2011) and Ansmann et al. (2012) estimated that the uncertainty in dust related backscatter coefficients is 15-20% in well-detected dessert dust layers and 20-30% in less pronounce aerosol layers. Moreover, we have calculated that the uncertainty of the dust occurrences presented in Sec. 3.1 ("% Dust / Used Overpasses"), might be up to 8% in latitudes away from the sources. Finally, an uncertainty induced in the dust product presented in this work, originates from the CALIPSO subtype selection algorithm. In this version of our product, both dust and polluted dust observations are considered polluted dust, and the pure dust component is separated using the dust separation method. The other aerosol layers, which are characterised as clean marine (CM), smoke (S), polluted continental (PC) or clean continental (CC) are considered to be cases clear of dust and are not tested for a dust component. This introduces negligible error in our analysis and is expected to induce a negative bias in the parameter "% Dust / Used Overpasses" less than 8%, mainly in areas above sea. In general, Clim-DE and Cond-DE products, the uncertainty of the dust extinction values close to the surface and at high latitudes is < 54%. At high altitudes and for latitudes up to 45°N, the uncertainty of the values is < 20%. Nevertheless, the standard deviation of the climatological products, coming from the natural variability of the dust events, may exceed to a large extent the uncertainty of the retrieval, reaching values as high as 100% and 200%.

In the latest release of CALIPSO Level 2 version 4 product (CALIPSO L2-V4, 2016), based on CALIPSO team announcement, the accuracy of the original CALIPSO product is increased and the uncertainty is reduced. This version is based on a revised calibration approach which leads to an increase in the total attenuated backscatter coefficients by ~3% overall as compared to the version 3 values (CALIPSO L1-V4, 2016). Several bugs are fixed and a major overhaul of the aerosol subtyping algorithms along with revisions on the lidar ratio selections is applied."

In Figure 1, there are some features that they look associated with the number of available observations, and consequently with the presence of clouds over the Mediterranean and Europe. I am not sure if the "%Dust/Used Overpasses" is enough to explain the DOD seasonal patterns in Europe. I would suggest to include an additional column with the number of used overpasses and to check how is working the algorithm of CALIOP to determine the corresponding aerosol subtype in this part of the domain.

[REPLY] The capability of CALIPSO to detect the dust subtype has been thoroughly evaluated by Burton et al., (2013) using a number of 109 underflights of CALIPSO with NASA's HSRL system and found that the detection of dust from CALIPSO is successful in 80% of the compared cases. Figure 1 is meant to present the frequency of CALIPSO dust occurences in the domain of our interest, in relation to the cloud-free overpasses of the sensor over the area. Following the helpful comment of the reviewer a new Table was added in the manuscript (Table 3), in order to provide a more informative representation of the dataset, including the percentages of the cloud free observations used, in relation to the total observations provided, aggregated on 6 areas over the study region. Furthermore we added the following discussion in the manuscript (Page 9, line 14):

Table 2 shows the impact of cloud contamination in our dataset. During AMJ, JAS and OND, more than 80% of the total observations are cloud-free above North Africa. Above Central-East Mediterranean (C-E Med.), more than 80% of the total observations are cloud-free and above Central West Mediterranean (C-W Med.) approximately 60% - 80% of the total observations are cloud-free. With increasing latitude, the cloud-free sampling is reduced to percentages of ~ 40% -60% in latitudes greater than 45° N. During JFM, cloudy conditions restrict our dataset in the greatest extent. During the same period, the cloud-free cases used represent ~ 80% of the total observations above North Africa, approximately 60 - 70% of the total observations above the Mediterranean and ~ 30% in the domain between 45° N - 60° N.

In the areas (and seasons) where clouds do not dominate (e.g. 70% clear-sky conditions), our cloud-free product is considered representative of the dust distribution. In areas where cloudy skies dominate (e.g. 30% clear-sky conditions), the clear-sky CALIPSO profiles cannot be considered as representative of all meteorological conditions, so the results should be used with caution."

In Page 10 Line 1, you mention that the results from Clim-DE can be used to estimate the impact of dust on cloud formation. As far as I understood, the EARLINEToptimized CALIPSO dust product is provided only for clear-sky conditions. In Sect. 4, you mention a recent paper from Mamouri and Ansmann (2016), but it is based on a ground-based lidar. Then, how could you estimate the dust impact on cloud formation from this EARLINET-optimized CALIPSO dust product?

[REPLY] The impact of dust on cloud formation is part of a second study we are working on. In this work, we use dust profiles from CALIPSO, in combination with EARLINET parameterizations, in order to calculate the dust mass concentration for particles with radius greater than 250nm and from there, based on known ice nuclei parameterizations to estimate ice nuclei concentration profiles. A detailed analysis of this technique is provided in the work of Mamouri and Ansmann (2016). An example of the application of this technique on collocated ground-based and CALIPSO data and comparison with in situ estimated ice nuclei will be presented in the upcoming ILRC Conference in Budapest. (Marinou, et al.: Lidar ice nuclei estimates and how they relate with airborne in-situ measurements, 28th ILRC, Bucharest, 25-30 July 2017). In order to keep the discussion as straightforward as possible, and to avoid confusing the readers, we decided to delete the corresponding part.

In my opinion, a further discussion about the similarities and discrepancies with other dust climatologies will enhance the impact of the results presented in the manuscript. Any comparison with other dust climatologies based on other datasets such as satellites (e.g. MODIS, AERONET, EARLINET or the official CALIPSO aerosol product); and models (as CAMS reanalysis or AEROCOM) is considered in the manuscript. Furthermore, how do the results of the present study improve those results of LIVAS (Amiridis et al., 2015)? These discussions will be useful for model evaluation, for example. Otherwise, it seems to me that some results are general and not enough justify in the manuscript.

[REPLY] We thank the reviewer for this comment. Amiridis et al., (2015), did not analyze the dust transport patterns. That paper describes the LIVAS database and focuses on the spectral conversion of the CALIPSO 532nm products for use in future ESA lidar missions that operate at 355nm. Following the suggestion of the reviewer, we calculated the optical depth using other available products such as the AOD from MODIS data and the DOD from MACC and RegCM4 models. We added a new figure, Figure 2, comparing our DOD seasonal maps with the ones produced with the above mentioned products, that is followed by a short discussion (page 10, line 26):

"In order to provide a more informative representation of the dust product presented here, we performed a comparison with MODIS AOD for the same period and the dust optical depth of the MACC reanalysis and a RegCM4 simulation for the period 2007-2012 and 2007-2014 respectively (Fig. 2). MODIS provides AOD for all natural and anthropogenic aerosol types. As a result the MODIS average value for the whole period and domain (0.267) is 281% almost three times, bigger than our product (0.095). It is noted thought that the values between the two satellite products are very similar over the Sahara desert. On the contrary, the corresponding average dust optical depth values of MACC (0.100) and RegCM4 simulations (0.104) reproduce better our product, since only dust is considered, though our product is lower by 5% and 8.6% respectively. Dust optical depth is overestimated over Europe and Mediterranean by MACC and RegCM4 simulations in comparison to our product in all seasons and especially in the hot periods AMJ and JJA, but the reasons of these discrepancies have to be further studied."

In Sect. 3.1 (Page 10) I don't understand the reason to include the dust mass concentration inversion results. This part of the discussion doesn't include any new insight with respect the analysis of the optical properties or any link to a particular previous study.

[REPLY] We removed the formulas and relevant discussion to the dust mass concentration inversion from the paper.

In Sect. 3.5, you could compare your results with a climatic index as the North Atlantic Oscillation Index (NAO) as Pey et at. (2013) did for PM10.

[REPLY] Following the analysis of Moulin et al. (1997) and Pey et al. (2013) we investigated the relation between NAO index and LIVAS AOD in seasonal and monthly basis (i.e., summer, winter, annually) for the period of our study but we did not find a statistically significant correlation and thus we did not include it in our results. Especially for summer the correlation between the NAO index and the LIVAS DOD over the western Mediterranean is negative but not statistically significant.

Is this de-seasonalised trend analysis sensitive to the number of available observations?

[REPLY] In our study, we calculate the DOD trend along with the statistical significance of each trend for the period 2007-2015 (108 monthly values). Nine years are considered a small period for a robust trend calculation and it would be interesting to repeat the same analysis in the future to an extended aerosol record. The de-seasonalization process as well as the trend are described only for the examined period only. In case we extend our analysis in the future by adding more years, results may change. We added this clarification in the manuscript (page 16, line 13):

"In our study, we calculate the DOD trend along with the statistically significance of each trend for the period 2007-2015 (108 monthly values). Nine years are considered a small period for a robust trend calculation and it would be interesting to repeat the same analysis in the future to extended aerosol record. The de-seasonalization process as well as the trend are describing the examined period only. Figure 7 shows the DOD internal variability of the 20 individual areas, as it is calculated from monthly mean DODs. Is evident from this figure that the DOD values in 2008 are relatively higher than the other years and in almost all the domains bellow 40°N. Similarly, relatively high values are observed in some of these areas for the year 2010. Since these years are at the beginning of our study period, they have a significant contribution on the negative trends observed during the examined period."

Minor comments:

Page 2 Line 13: Add Nickovic et al. (2016).

[REPLY] We added this reference.

Page 2 Line 16: Add Granados-Muñoz et al. (2016) and Bovchaliuk et al. (2016).

[REPLY] We added these references.

Page 3 Line 25: Replace Gkikas et al. (2015, ACPD) by Gkikas et al. (2016, ACP).

[REPLY] It is replaced.

Page 3 Line 30: When you said "large scale statistics of discriminate and optimized dust extinction and AOD fields from CALIPSO", what does it mean? What about Amiridis et al. (2013) and Amiridis et al. (2015)?

[REPLY] This phrase is removed and replaced by the phrase (page 4, line 4):

"To our knowledge, this is the first time that a 3D pure-dust dataset is statistically analysed over the area of North Africa and Europe in order to provide not only the horizontal but also the vertical patterns of Saharan dust intrusion in the Mediterranean."

Amiridis et al., (2013 and 2015) did not analyze the dust transport patterns. Amiridis et al (2013) describe the methodology for the pure-dust retrieval algorithm. Amiridis et al., (2015) describes the LIVAS database. The later paper focuses on the spectral conversion of the CALIPSO 532nm products for use in future ESA lidar missions that operate at 355 nm.

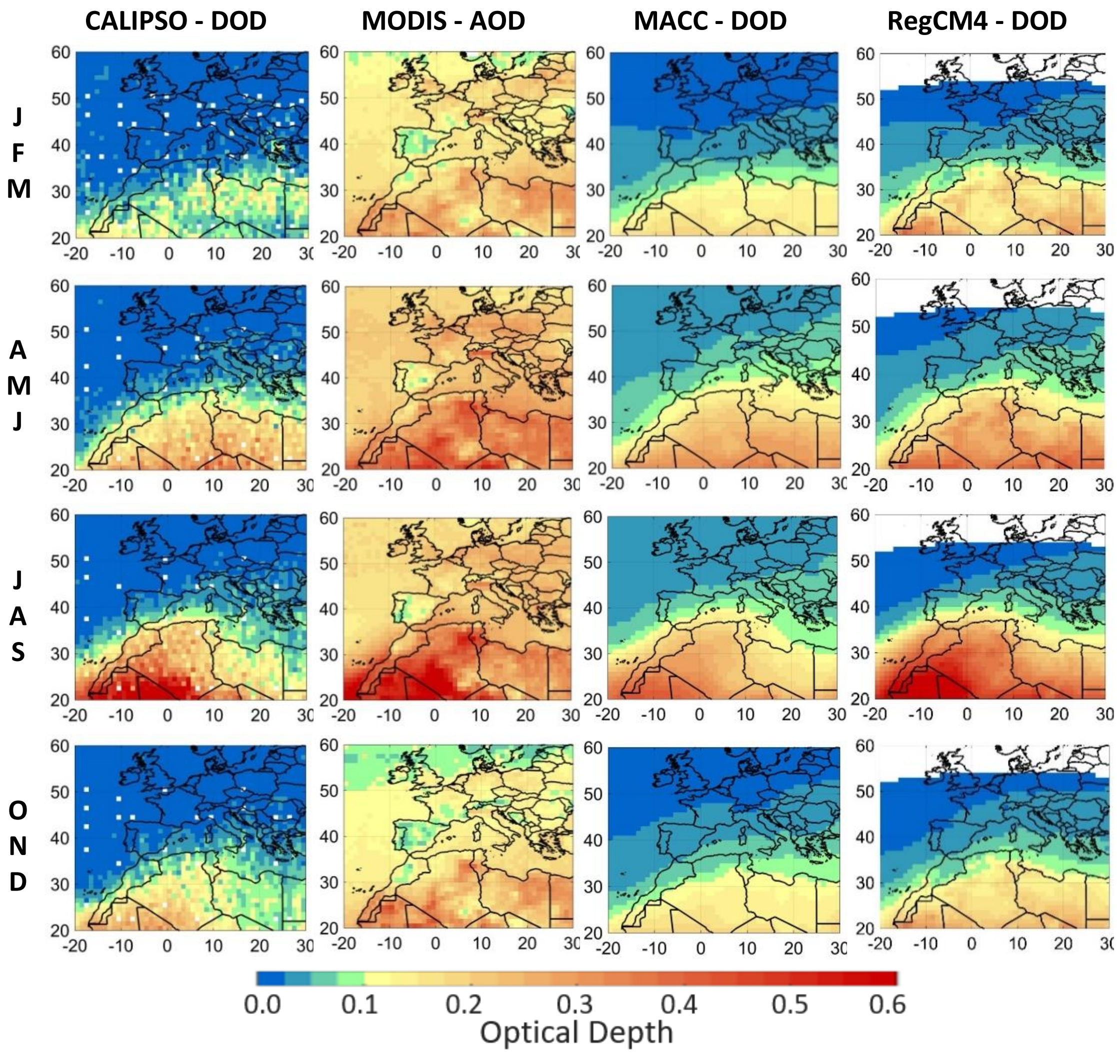
Sect. 3.1: It would be good if you can add a short comparison of the resulting DOD seasonal maps with the results of MODIS, MISR or any available reanalysis (as CAMS or MERRA).

[REPLY] We thank the reviewer for this comment. Therefore we calculated the optical depth using other available products such as the AOD from MODIS data and the DOD from MACC and RegCM4 models. We added a new figure, Figure 2, comparing our DOD seasonal maps with the ones produced with the above mentioned products, that is followed by a short discussion (page 10, line 26):

"In order to provide a more informative representation of the dust product presented here, we performed a comparison with MODIS AOD for the same period and the dust optical depth of the MACC reanalysis and a RegCM4 simulation for the period 2007-2012 and 2007-2014 respectively (Fig. 2). MODIS provides AOD for all natural and anthropogenic aerosol types. As a result the MODIS average value for the whole period and domain (0.267) is 281% almost three times, bigger than our product (0.095). It is noted thought that the values between the two satellite products are very similar over the Sahara desert. On the contrary, the corresponding average dust optical depth values of MACC (0.100) and RegCM4 simulations (0.104) reproduce better our product, since only dust is considered, though our product is lower by 5% and 8.6% respectively. Dust optical depth is overestimated over Europe and

Mediterranean by MACC and RegCM4 simulations in comparison to our product in all seasons and especially in the hot periods AMJ and JJA, but the reasons of these discrepancies have to be further studied."

Figure 2: Comparison of the seasonal spatial distribution of the optical depth as received by (first column) pure-dust CALIPSO DOD product, (second column) MODIS AOD product, (third column) MACC reanalysis DOD product, (fourth column) RegCM4 simulated DOD product.



Sect. 3.2: In this case, you could compare your results from those obtained from EARLINET or models.

[REPLY] We thank the reviewer for this suggestion. We added the following paragraph in section 3.2 (page 11, line 29):

"In general, our results are in agreement with lidar-based studies which have been performed in several European sites. Papayannis et al. (2008) performed an exhaustive analysis on Saharan dust particles over Europe using EARLINET lidar profiles. They found that the dust layer center of mass extends from 3.0 to 3.8 km and the thickness ranges from 0.7 to 3.4 km. Specifically, Balis et al. (2012) calculated the mean base and top of dust layers in the eastern Mediterranean, Thessaloniki, to be around 2.5 ± 0.9 km and 4.2 ± 1.5 km, respectively. More recently, Mona et al. (2014) analyzed a long dataset of Saharan dust intrusions over Potenza, Italy, and found a mean layer centre of mass of 3.5 ± 1.5 km."

Sect 3.5: how do your results fit with those showed in Gkikas et al. (2016)?

[REPLY] In Gkikas et al. (2016) there is no discussion on the interannual variability of the dust events. The interannual variability of dust events is discussed in Gkikas et al. (2013). A sentence is added in the manuscript comparing our results to this work (page 15, line 32):

"Furthermore, over the same domain the decreasing trend of DOD coincides with the decrease of Saharan desert dust episodes as reported by Gkikas et al. (2013)."

Page 10 Line 28: Add Huneeus et al. (2016).

[REPLY] We added this reference.

Page 10 Line 30: "it is likely that the surface and elevated dust have different origins" sounds speculative. You could check this assumption with models or back trajectories.

[REPLY] The corresponding sentence is removed from the revised manuscript.

Figure 3. I would use the same colour palette than in Figure 4.

[REPLY] Figures 3 and 4 (new Figures 4 and 5) are based on the same color pallet with the difference that in Figure 4 a new restriction is introduced. In both cases, the black values represent the mean terrain elevation. In Figure 5, when we average the Con-DE, there is different sampling than the one used for the Clim-DE, and as a result, some means are produced from very few numbers of dusty observations (dO). In order to filter these case from the plot, so as the readers to concentrate on more significant features, we mask them with the new gray color. In order to better address these filters in the plots we changed the color pallet of Figure 5, and added a NaN black box (similar as in Figure 4), and labeled the gray box as "<4dO".

We also changed the manuscript when introducing the two figures:

In section 3.3 (page 12, line 11): "The median surface elevation is depicted with black colour (and is labeled as NaN) in the plots.

In section 3.4 (page 14, line 3): "Con-DE values derived from less than 4 dust observations (dO) in each cell are masked with grey colour (and are labeled as <4dO) in the plots. The median surface elevation is depicted with black colour (same as in Fig. 4)."

Figure 3. Could you provide any further explanation about the sharp transition over the Atlas?

[REPLY] We added the following sentence in the paper (page 12, line 32):

"The area south of Atlas Mountains (Fig. 4e, f, g, h) is characterized by haboob activity (Knippertz et al., 2009; Solomos et al., 2012). These systems are generated from convective outflows and contribute to the interannual burden of dust at this area."

Figures 3,4,5: Longitude and Latitude labels can be removed. They are too small.

[REPLY] The labels size is increased and it is more visible in the new version of the manuscript.