Anonymous Referee #1

We would like to thank the anonymous referee for providing a series of comments and suggestions that helped us to improve the submitting paper. In the following, answers to comments are reported in italics, just below each related comment. When needed, the part of the manuscript we modified or added to the old version is reported in bold.

GENERAL COMMENTS:

<u>EARLINET has currently 27 active stations and the contributing stations have been performing correlative measurements with CALIPSO satellite, would be interesting taking into account data from other stations beside the five one used. It would increase the number of comparable data, reducing the uncertainties from spatial and temporal differences. One important conclusion for the differences observed on CALIPSO and ground-based retrievals is the difference in sampling volumes and the spatial variability of the aerosol fields, which is expected when validating satellite data. However, the investigation of the influences of air masses trajectories between ground-based lidar stations and CALIPSO overpasses region should be considered in order to reduce or at least justify these differences.</u>

Two important points are correctly underlined by the Referee #1: the number of ground-based data used for the CALIPSO data investigation and the influence of air masses on the comparison.

About the first point, it is important to highlight the main aim of this paper: the investigation of the reliability and significance of CALIPSO climatological data. Keeping this in mind, the methodology described in Section 2.3 was adopted for the construction of the EARLINET dataset for this study. The main concept is to consider only site-specific datasets with good coverage of monthly profiles resulting from simultaneous CALIPSO overpasses within 100 km horizontal distance.

Apart from the data used and reported in the manuscript (Table 3, Page 31232), other profiles in correspondence with CALIPSO overpasses are available from additional seven stations as listed in the following table:

Table: EARLINET observations for the stations not included in the analysis.

Station	*Observations*	*Monthly profiles*
Athens (GR)	*21*	*1*
Barcelona (E)	*12*	*1*
Bucharest (RO)	*2*	*-*
Cabauw (NL)	*15*	*2*
Madrid (E)	*31*	*4*
Maisach (D)	*10*	*1*
Thes/niki (GR)	*8*	*-*

For six of them the number of monthly profiles is lower than 3 for the 2006-2010 period. For the Madrid station, 4 monthly profiles are available which is still a low number but could be considered in the dataset. However, the Madrid profiles relevant for this study are provided with a coarse range resolution (about 400 m). This does not allow the investigation of each aerosol layer which is the cornerstone of the analysis reported in the paper. This analysis is based indeed on a fine vertical *resolution as is needed to identify with high confidence the geometrical characteristics of each aerosol layer. Regarding this remark, the following phrase has been implemented in Section 2.3.4:*

"Apart from the data redundancy, the stations were also selected with respect to their range resolution. The analysis is based on the precise layer location, which can be accomplished by using a resolution finer or comparable to CALIPSO one (60 m in the lower troposphere)."

For what concerns the difference in sampling volumes and the spatial variability of the aerosol fields the referee is right, these aspects should be carefully considered in this kind of study. In that sense, we selected only EARLINET correlative measurements. Limiting ourselves to this dataset strongly decreases the number of data available for the analysis (Page 31205 – lines 10-11), but minimizes the spatial variability. The problem of sampling error and spatio-temporal variability in the EARLINET-CALIPSO comparison was considered already at the time of the planning of EARLINET measurement for CALIPSO validation purposes. The impact of the spatio-temporal distance on EARLINET-CALIPSO comparison was investigated for different stations in devoted papers (e.g., Mona et al., 2009; Mamouri et al., 2009). At network level we found that for distance below 100 km the discrepancies in the signal (CALIPSO Level 1 data) are below 5%. Moreover, for cases of long-range transported aerosol like Saharan dust, it was found that a horizontal distance of 100 km corresponds to high correlation among the two profiles (Pappalardo et al., 2010). In Section 2.2 we added the next lines.

"In this kind of measurements, the atmospheric variability both in time and space is a fundamental point. The impact of the spatio-temporal distance on EARLINET-CALIPSO comparison was investigated for different stations in devoted papers (e.g., Mona et al., 2009; Mamouri et al., 2009). At network level we found that for distance below 100 km the discrepancies in the signal (CALIPSO Level 1 data) are below 5%. Moreover, for cases of long range transported aerosol like Saharan dust, it was found that a horizontal distance of 100 km corresponds to high correlation among the 2 profiles (Pappalardo et al., 2010)."

For the sake of completeness, HYSPLIT (Draxler and Hess, 1998) model in backward mode was used to check the air masses movement and if EARLINET and CALIPSO simultaneous measurements sampled the same air volumes for all the cases used in this paper. The model was initiated for each CALIPSO measurement and its EARLINET counterpart and the corresponding trajectories were visually inspected. Each model run was set in the range of 0.5-6 km a.s.l. and for constant height levels, independently of the existence of aerosol layers. For all the cases related to this study, the model analyses indicated that the ground based and satellite lidars sampled the same air mass. We inserted the text below in Section 2.3.4.

"To ensure that the same air volumes were sampled, HYSPLIT model (Draxler and Hess, 1998) in backward mode was used. The model was initiated for each CALIPSO measurement and its EARLINET counterpart and the corresponding trajectories were visually inspected. Each model run was set in the range of 0.5-6 km a.s.l. and for constant height increments, independently of the existence of aerosol layers. For all the cases related to this study, the model analyses indicated that the ground based and satellite lidars sampled the same air mass."

GENERAL QUESTIONS AND COMMENTS FOR CONSIDERATION:

Subsection 2.3.1 – page 31205: It is described the comparison methodology between CALIPSO CL3 products and EARLINET retrievals. Please, could the authors explain in more details how the CL3* products were produced?

The spatio-temporal discrepancies of the EARLINET and CL3 datasets, as explained in Page 31205 (lines 3-11), require that the comparison is limited only to simultaneous CALIPSO and EARLINET observations. Therefore, starting from the available EARLINET profiles in correspondence to CALIPSO overpasses, we obtain CL3 profiles following the steps as listed below:*

1. *We select the CALIPSO Level 2 data found within a 2°x2° grid that contains each EARLINET site (Page 31236 – Figure 2).*
2. *We screen the CALIPSO data, following the rubric described in Winker et al. (2013). Although the following condition is modified: Extinction_Coefficient_Uncertainty_532≤10 km^{-1}.*
3. *We exclude samples where the screening criteria are invoked. Moreover, samples that represent clear air are assigned a value of 0.0 km^{-1}. Then, the mean profile is retrieved.*
4. *We average mean profiles obtained following the above steps within the same month, obtaining a CL3* profile.*

In conjunction with comments made from Referee #2, the second paragraph in Section 2.3.1 (Page 31205 – lines 12-26 and Page 31206 – lines 1-6) with respect to the production of CL3 data has been changed in the revised version of the manuscript and now reads like:*

"To produce the CL3* monthly profiles, we use the CL2 Version 3.01 Aerosol Profile product, which includes aerosol extinction and backscatter coefficient profiles at 532 nm. The spatial domain onto which the CL2 data are mapped is nearly 2°×2° and contains the EARLINET sites. This means that the longitudinal resolution is smaller owing to the distance of CALIPSO overpasses (<100km) from the EARLINET measuring site. The 6-step methodology to quality assure the CL3 profiles (Winker et al., 2013; Appendix A) is modified adjusting an existing metric according to the rubric used by Campbell et al. (2012). In particular, the metric is adjusted as:

Extinction_Coefficient_Uncertainty_532≤10 km^{-1}.

The lower boundary, here, is set to a smaller value, whereas within CALIPSO procedure, retrievals deemed unstable are set to 99.9 km^{-1}. In this case, samples that meet this condition are removed as well as samples at lower altitudes. Prior to averaging, samples are excluded where the screening criteria are invoked and moreover, for samples that represent clear air a value of 0.0 km^{-1} is assigned. Although, clear air samples over the surface are ignored from the averaging process in the case that the base of the lowest aerosol layer in the profile is below 2.5 km."

Page 31206, line 9: Can you consider two measurements representative of a month? For cases with only two lidar measurements, how many CALIPSO measurements were used to produce CL3* products?

Indeed, two measurements cannot be considered representative for one month. This is the reason why we do not use the original CL3 data: on average we would have seven nighttime CALIPSO profiles averaged to be compared against EARLINET monthly profile obtained from 2-3 files (7 is the mean number of nighttime observations for the five sites and the period 2006-2010). On the other hand, the CL3 product and EARLINET monthly averages, include exactly the same number of profiles. Each CALIPSO profile was compared to its EARLINET counterpart, eliminating in this way any temporal* *discrepancies. For example, if during one month two ground-based lidar measurements are available, two CALIPSO profiles are used for calculating the monthly average to be compared with.*

Page 31206, line 28: Is the term approximate particle depolarization ratio or volume depolarization ratio?

According to Omar et al. (2009) the term used in the CALIPSO aerosol typing scheme is the corrected depolarization ratio (or estimated particle depolarization ratio) and is denoted as δ_V (Eq. 10 of Omar et al., 2009). However, the term approximate (or approximated) particle depolarization ratio is used by various studies on CALIPSO products evaluation (e.g., Amiridis et al., 2013; Burton et al., 2013; Tesche et al., 2013). In conformity with this evaluation and recent studies, we adopted the "approximate particle depolarization ratio" as nomenclature.

Page 31211, line 1: What would be the causes for the discrepancies between extinction and backscatter profiles in the lowermost part of the profile between CALIPSO and Granada station (figures 3b and 4b)?

The main element of this discrepancy is the complex topography of the station (Alados-Arboledas et al., 2003; Navas-Guzman et al., 2013). The mean CALIPSO ground-track distance from the ground-based lidar is 66.8 km (Page 31232 – Table 3) and ensures the sampling of the same air volumes. However, the aerosol content is likely to differ between the ground-based lidar and the CALIPSO observations as the mountains around the EARLINET station could act as a physical barrier: anthropogenic pollution or low-lying dust plumes could be blocked either way. To this direction, the typing comparison (see Page 31241 – Figure 7) showed that anthropogenic particles were not identified from CALIPSO while for EARLINET these particles are dominant in the lowermost part of the ground-based profiles. The next phrase is inserted in the manuscript.

"The backscatter comparison (Fig. 4b) revealed the same characteristics with enhanced discrepancy in the lowermost part of the profile, as expected due to the complex topography of the region (Guerrero-Rascado et al., 2008)."

figure 5b-page 31239: In the lidar ratio profile for Granada station is presented the lidar ratio signal starting at 2 km approximately. How is the procedure to classify or identify the aerosol subtype in the region between 1 - 2 km presented in figure 7a, since the lidar ratio signal is missing in this region?

This should be clarified in the text. The 1-2 km height interval refers only to range 1.7-2.0 km, which is the range where EARLINET yields values. Regarding the CALIPSO bar plot (see Figure 7b) the height interval 1-2 km is the same as for EARLINET in order to compare same portions of the height. The next phrase will be inserted in Section 3.

"However, the first bin is associated with the lowest altitude point retrieved by EARLINET, thus the range can be smaller than 1 km. For this comparison, the same distance was used for both EARLINET and CALIPSO typing. For the sake of visual consistency, the height bins are kept equidistant for all the plots."

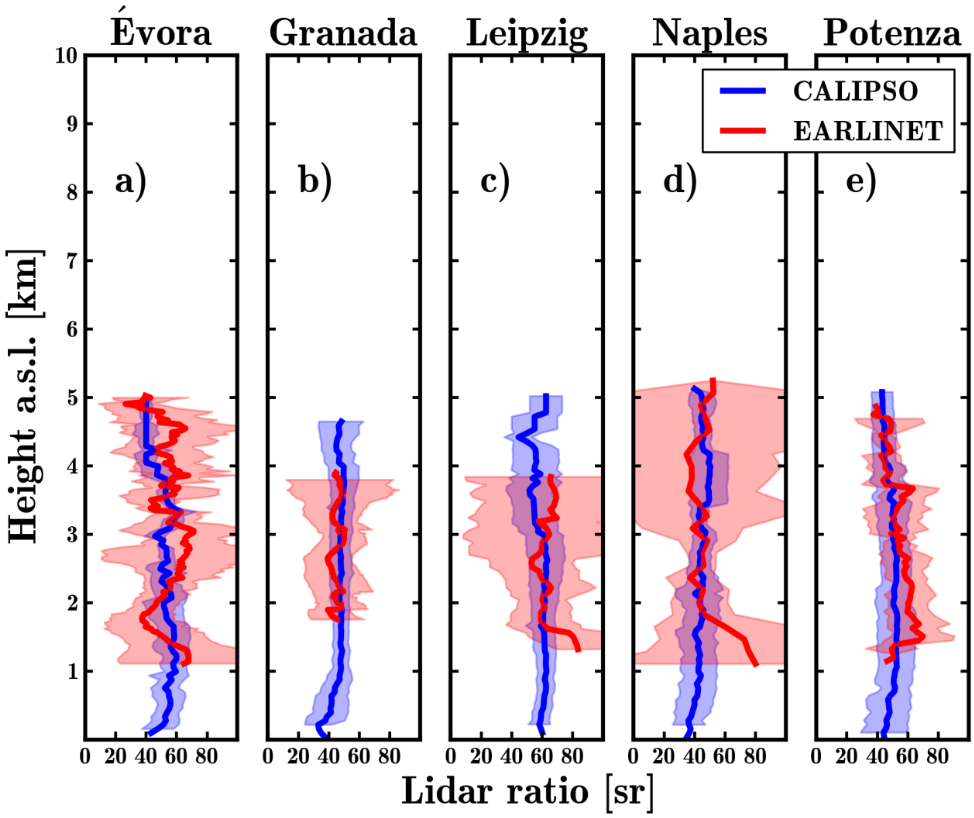
Page 31212, line 13: "The CALIPSO typing, shown in Fig. 8b, for the height interval 1–2km identifies Smoke and Polluted Continental equally". If the CALIPSO algorithm uses the layer altitude to classify the aerosol between Smoke or Polluted continental, I'm wondering, why there are aerosol layers between 1 - 2 km classified as smoke over Leipzig station?

It is true that this point needs further clarification. According to CALIPSO typing scheme, only elevated layers can be classified as smoke particles (Omar et al., 2009; see Figure 2), suggesting that smoke layers cannot be in contact with the surface of the Earth. The algorithm follows between two pathways (pathways 7 and 9, Figure 2 of Omar et al., 2009) in order to discriminate smoke and polluted continental samples. The attribute that defines this selection is whether the sample is elevated, even if at very low altitudes. For Leipzig CL3 data, smoke plumes were found to lie as low as ~0.5 km a.s.l. whereas Polluted Continental extended from the ground to higher altitudes. Specifically, for the range 1-2 km Polluted Dust, Dust, Clean Continental, Polluted Continental, and Smoke particles were present and accordingly to CALIPSO typing scheme are aerosol types that can be observed over land (pathways 3-7 and 10, Figure 2 of Omar et al., 2009). Forest-fire smoke particles can be due to long-transported plumes either from North America or rarely from Siberia (e.g., Mattis et al., 2008).*

Page 31213, line 10: "In the region of 3 - 4 km there is good agreement between the two platforms with mean lidar ratio values of S_{aer}= 44±4srfor Naples station and S_{aer}= 44±2srfor CALIPSO". However, in figure 5d is missing the lidar ratio profile between 3 - 4 km for Naples station. Would be this agreement of S_{aer} = 44 sr in the region of 2 - 3 km? Why is the profile missing between 3 - 4 km? How can this missing lidar ratio information can compromise the confidence of the EARLINET aerosol typing between 3 - 4 km presented in figure 9a?

The range discussed in Page 31213 – line 10 is wrong. The confusion regarding the missing part of the profile is due to our mistake. The text and figure have been corrected in the revised version of the manuscript.

"In the region of 2-3 km there is good agreement between the two platforms with mean lidar ratio values of S_{aer} = 44±4 sr for Naples station and S_{aer}= 44±2 sr for CALIPSO."



Page 31239 – Figure 5: Lidar ratio at 532 nm for CL3* (blue line) and for EARLINET (red line). From left to right: (a) Evora, (b) Granada, (c) Leipzig, (d) Naples, and (e) Potenza.

Page 31214, line 9: "The lower level disparity typically is weakened during summer months, and it is intensified in winter, yet the sample size is too small to quantify the periodicity of this discrepancy". Despite the difficult to obtain a large quantity of coincident data between CALIPSO and ground-based lidars, would be interesting to mention what is the period/season of the year the most of data were obtained and what kind of discrepancies or influences can produced in this validation study.

The referee is correct that the seasonal comparison would be of high interest. Ground-based lidar measurements are limited in presence of low-lying thick clouds and during precipitation. Thus, most of the measurements were made during summer and spring as reported in the following table. This means that the analyzed dataset is highly influenced by long-range transported dust/smoke particles as more than 80% of the collected profiles correspond to months favoring this aerosol situation. Clean conditions are less represented in these datasets, but on the other hand these cases are also less significant in terms of AOD (Mona et al., 2012). The influence of lidar ratio increases with increasing layer AOD. Therefore, even if the data correspond greatly to warm months, we assume that on the findings regarding the CALIPSO typing and lidar ratio impact the situation will not alter significantly. The table and the following phrase are inserted in Section 2.2.

Table: Seasonality of the available monthly profiles.

Season	Monthly profiles
Summer	25
Spring	13
Autumn	8
Winter	1

"Over 80% of the observations were performed during summer and spring months owing to the favorable weather conditions and do not permit to assess the seasonal behavior of the comparison."

Moreover, the next paragraph will be reported in the Conclusions:

"The larger number of available comparisons for the warmer months, indeed, influences in some way our results. The analyzed dataset is highly influenced by dust/smoke presence which typically occurs during these months. Clean conditions are less represented, here, but on the other hand these cases have also less impact in terms of AOD. However it should be noted that the influence of lidar ratio increases with the layer AOD so it is more relevant for the dust/smoke plumes in general. Therefore, even if the data correspond greatly to warm months, we assume that on the findings regarding the CALIPSO typing and lidar ratio assessment the situation will not alter significantly."

Page 31215, line 6: Why the relative differences of the extinction and backscatter comparison presented in figure 11 are so large for elevated altitudes? How the mean relative differences were calculated, EARLINET-CALIOP/EARLINET?

The differences are calculated as $(x_{CALIPSO} - x_{EARLINET})/x_{EARLINET}$, where x is either the backscatter or the extinction profile. At high altitudes the relative difference yields high biases because the ratio consists of very small numbers. This comment is also in agreement with comment #7 from Referee #2 and now the relative difference is treated differently and clearly explained in the text. The comparison between extinction and backscatter relative difference is now reported only for altitude below which the 90% of the columnar AOD is confined, as suggested by Referee #2. Discussion relative to the figure was correspondently modified.

Page 31216, line23: Would be interesting to present values of marine lidar ratio retrieved by the EARLINET stations for cases of mixture, in order to check the disagreements between the lidar ratio values assigned by CALIPSO. It can help to improve the CALIPSO algorithm for polluted dust aerosol subtype, for instance.

For the plots in Pages 31240-31244, the Marine subtype for the EARLINET typing unequivocally refers to clean marine plus marine mixtures (Page 31207 – lines 20-23). However, in Section 3.3 the maritime particles mixtures were omitted in order to ensure simultaneous subtype identification by EARLINET and CALIPSO. CALIPSO subtypes do not include mixed marine layers. Following the referee's comment we included an extra line in Table 5 (Page 31234) and the next phrase was added in the Section 3.3.

“This study, also, estimated a mean lidar ratio for mixed marine particles of 33±5 sr, which is consistent with values reported in literature (Müller et al., 2007; Gross et al., 2011; Burton et al., 2013). CALIPSO typing scheme does not incorporate marine mixtures in a separate subtype, therefore a comparison is not feasible.”

Page 31234 – Table 5: Mean lidar ratio at 532 nm for the various aerosol subcategories as measured by EARLINET sites and corresponding statistical parameters. The last column refers to the CALIPSO lidar ratio assumed values and their associated lidar ratio distributions (mean plus standard deviation). M stands for Marine, D for Dust, PC for Polluted Continental, CC for Clean Continental, PD for Polluted Dust, S for Smoke, and MM for Mixed Marine subtype. Note that, here, the M subtype corresponds to pure marine particles.

	EARLINET				CALIPSO
Aerosol type	Mean±SD [sr]	Range [sr]	Median [sr]	# Samples	Mean±SD [sr]
M	23±3	21-24	22	5	20±6
MM	33±5	25-38	34	8	-
D	51±10	41-73	48	16	40±20
PC	62±10	51-78	61	14	70±25
CC	47±4	44-52	46	4	35±16
PD	53±14	35-78	49	13	55±22
S	67±10	54-80	65	11	70±28

REFERENCES

Alados-Arboledas, L., Alados, I., Foyo-Moreno, I., Olmo, F., and Alcántara, A.: The influence of clouds on surface UV erythemal irradiance, Atmos. Res., 66, 273–290, 2003.

Amiridis, V., Wandinger, U., Marinou, E., Giannakaki, E., Tsekeri, A., Basart, S., Kazadzis, S., Gkikas, A., Taylor, M., Baldasano, J., and Ansmann, A.: Optimizing CALIPSO Saharan dust retrievals, Atmos. Chem. Phys., 13, 12089–12106, doi:10.5194/acp-13-12089-2013, 2013.

Burton, S. P., Ferrare, R. A., Vaughan, M. A., Omar, A. H., Rogers, R. R., Hostetler, C. A., and Hair, J. W.: Aerosol classification from airborne HSRL and comparisons with the CALIPSO vertical feature mask, Atmos. Meas. Tech., 6, 1397–1412, doi:10.5194/amt-6-1397-2013, 2013.

Campbell, J. R., Tackett, J. L., Reid, J. S., Zhang, J., Curtis, C. A., Hyer, E. J., Ses-sions, W. R., Westphal, D. L., Prospero, J. M., Welton, E. J., Omar, A. H., Vaughan, M. A., and Winker, D. M.: Evaluating nighttime CALIOP 0.532 μm aerosol optical depth and extinction coefficient retrievals, Atmos. Meas. Tech., 5, 2143–2160, doi:10.5194/amt-5-2143-2012, 2012.

Draxler, R.R., and Hess, G.D.: An overview of the HYSPLIT_4 modeling system of trajectories, dispersion, and deposition. Aust. Meteor. Mag., 47, 295-308, 1998.

Groß, S., Tesche, M., Freudenthaler, V., Toledano, C., Wiegner, M., Ansmann, A., Althausen, D., and Seefeldner, M.: Characterization of Saharan dust, marine aerosols and mixtures of biomass burning aerosols and dust by means of multi-wavelength depolarization-and Raman-measurements during SAMUM-2, Tellus B, 63, 706–724, 2011.

Guerrero-Rascado, J.L., Ruiz, B., and Alados-Arboledas, L.: Multi-spectral Lidar characterization of the vertical structure of Saharan dust aerosol over southern Spain, Atm. Env., 42, 2668-2681, http://dx.doi.org/10.1016/j.atmosenv.2007.12.062., 2008.

Mamouri, R. E., Amiridis, V., Papayannis, A., Giannakaki, E., Tsaknakis, G., and Balis, D. S.: Validation of CALIPSO space-borne-derived attenuated backscatter coefficient profiles using a ground-based lidar in Athens, Greece, Atmos. Meas. Tech., 2, 513–522, doi:10.5194/amt-2-513-2009, 2009.

Mattis, I., Müller, D., Ansmann, A., Wandinger, U., Preißler, J., Seifert, P., and Tesche, M.: Ten years of multiwavelength Raman lidar observations of free-tropospheric aerosol layers over central Europe: geometrical properties and annual cycle, J. Geophys. Res.-Atmos., 113, d20202, doi:10.1029/2007JD009636, 2008.

Mona, L., Pappalardo, G., Amodeo, A., D'Amico, G., Madonna, F., Boselli, A., Giunta, A., Russo, F., and Cuomo, V.: One year of CNR-IMAA multi-wavelength Raman lidar measurements in coincidence with CALIPSO overpasses: Level 1 products comparison, Atmos. Chem. Phys., 9, 7213–7228, doi:10.5194/acp-9-7213-2009, 2009.

Mona, L., Liu, Z., Müller, D., Omar, A., Papayannis, A., Pappalardo, G., Sugimoto, N., Vaughan, M.: Lidar Measurements for Desert Dust Characterization: An Overview, Advances in Meteorology, vol. 2012, pp. 36, doi:10.1155/2012/356265, 2012.

Müller, D., Ansmann, A., Mattis, I., Tesche, M., Wandinger, U., Althausen, D., and Pisani, G.: Aerosol-type-dependent lidar ratio observed with Raman lidar, J. Geophys. Res., 112, D16202, doi:10.1029/2006JD008292, 2007.

Navas-Guzmán, F., Bravo-Aranda, J., Guerrero-Rascado, J., Granados-Munoz, M., and Alados-Arboledas, L.: Statistical analysis of aerosol optical properties retrieved by Raman lidar over Southeastern Spain, Tellus B, 65, 21234, doi:10.3402/tellusb.v65i0.21234, 2013.

Omar, A., Winker, D., Kittaka, C., Vaughan, M., Liu, Z., Hu, Y. X., Trepte, C., Rogers, R., Ferrare, R., Lee, K., Kuehn, R., and Hostetler, C.: The CALIPSO automated aerosol classification and lidar ratio selection algorithm, J. Atmos. Ocean. Tech., 26, 1994–2014, doi:10.1175/2009jtecha1231.1, 2009.

Pappalardo, G., Wandinger, U., Mona, L., Hiebsch, A., Mattis, I., Amodeo, A., Ansmann, A., Seifert, P., Linné, H., Apituley, A., Alados Arboledas, L., Balis, D., Chaikovsky, A., D'Amico, G., De Tomasi, F., Freudenthaler, V., Giannakaki, E., Giunta, A., Grigorov, I., Iarlori, M., Madonna, F., Mamouri, R., Nasti, L., Papayannis, A., Pietruczuk, A., Pujadas, M., Rizi, V., Rocadenbosch, F., Russo, F., Schnell, F., Spinelli, N., Wang, X., and Wiegner, M.: EARLINET correlative measurements for CALIPSO: first intercomparison results, J. Geophys. Res., 115, D00H19, doi:10.1029/2009JD012147, 2010.

Tesche, M., U. Wandinger, A. Ansmann, D. Althausen, D. Müller, and A. H. Omar: Ground-based validation of CALIPSO observations of dust and smoke in the Cape Verde region, J. Geophys. Res. Atmos. , 118, 2889–2902, doi:10.1002/jgrd.50248, 2013.

Winker, D. M., Tackett, J. L., Getzewich, B. J., Liu, Z., Vaughan, M. A., and Rogers, R. R.: The global 3-D distribution of tropospheric aerosols as characterized by CALIOP, Atmos. Chem. Phys., 13, 3345–3361, doi:10.5194/acp-13-3345-2013, 2013.