We appreciate the Referees comments that helped us to improve the submitting paper. In the following, answers to comments are reported in italics, just below each related comment. The comments are underlined and numbered to ease the process. When needed, the part of the manuscript we modified or added to the old version is reported in bold. Moreover, the references of the cited literature are given in the end of the document.

Referee Comment #1 (S. P. Burton)

Specific comments:

1. Figures 2 and 7. What are the resolution of the extinction and backscatter profiles? How is the lidar ratio calculated? Were the extinction and backscatter at the same resolution before taking the ratio? I ask because there are differences in the shape of extinction and backscatter in each of the discussed layers that do not seem particularly consistent with the idea that each layer is a specific coherent aerosol type. For example, the "layer" below 2 km in Figure 2 has a completely different shape in backscatter vs. extinction, leading to large variability in the lidar ratio, much larger than the suggested error bars. Do you think this variability is real, or could it be that the local maximum (seen in backscatter) is smoothed out in extinction by a coarser vertical resolution? If it's real, is it likely this is a single consistent aerosol type in this layer? Similarly, the different slopes in the "layer" between 3.5 and 5 km lead to a very large slope in the lidar ratio that does not seem consistent with the idea that this is a single aerosol type. If this variability is spurious, it is liable to create additional apparent noise in the classification that is not really related to aerosol variability within classes. (Of course, spurious error would also be a concern in general, not just for classification.)

P34-Figure 2 and P39-Figure 7 in the submitted paper are reporting the profiles at their highest resolutions, i.e. the particle backscatter coefficient profiles have a higher resolution when compared to the extinction ones and typically the ultraviolet profiles have a better resolution than the visible profiles. The lidar ratio profiles have the same resolution of the particle extinction coefficient profiles. Prior to the calculation of the lidar ratio, the particle backscatter coefficient profiles are smoothed using a 2nd order Savitzky-Golay filter at an effective vertical resolution that varies with height, for more details see Iarlori et al. (2015).

In particular, for P34-Figure 2, the resolution of the particle backscatter coefficient profiles is 60 m while for the particle extinction coefficient the resolution varies with height. The effective resolution for 355 nm (blue line) and 532 nm (red line) is given in Figure A and follows the procedure described in Pappalardo et al. (2004). For the height range 1.5-2.5 km, the highest resolution is 120 m for 355 nm while for 532 nm varies from 120 m to 480 m. Higher in the atmosphere, the resolution degrades with height (faster for 532 nm) and becomes constant at 4.5 km for 355 nm (3.5 km for 532 nm).

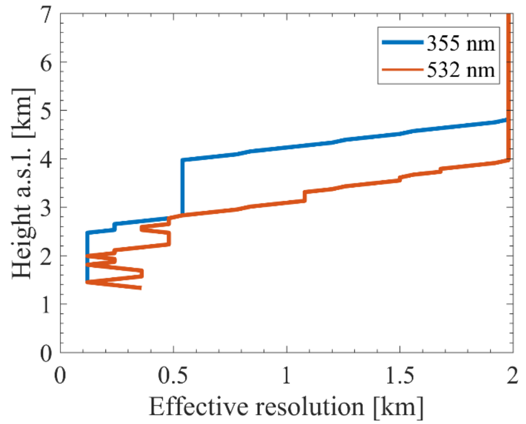

Figure A: *The effective resolution of the extinction coefficient at 355 nm (blue line) and 532 nm (red line) for the MUSA system in Potenza on 14/07/2011, 19:20–22:10 UTC.*

Figure B shows the same panels as for P34-Figure 2 along with the smoothed backscatter profiles (panel b) that were used for the calculation of the lidar ratio. The next lines are implemented in the revised manuscript:

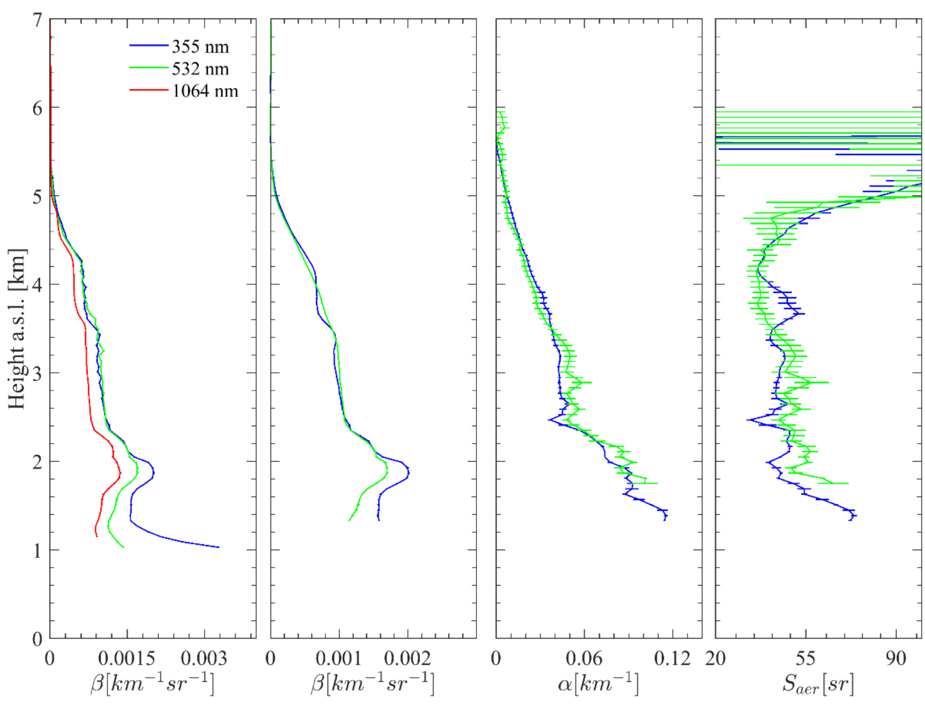

Figure B: *Optical profiles measured in Potenza, on 14 July 2011, 19:20–22:10 UTC with a multiwavelength Raman lidar. From left to right, (a) the particle backscatter coefficient with the full resolution, (b) the smoothed particle backscatter coefficient, (c) the particle extinction coefficient, and (d) the particle lidar ratio. The error bars correspond to the standard deviation.*

The particle extinction and backscatter coefficient are given with their full resolution. To calculate the lidar ratio, the backscatter coefficient was smoothed in the same effective vertical resolution using a Savitzky-Golay second order filter (Iarlori et al., 2015) and only the useful range of signals was kept; the effective resolution of the resulting profiles varied from 120 m to 480 m using the method described in Pappalardo et al. (2004).

About the height-independent shape of the intensive properties, in the revised version of the paper we rephrased the sentence because it is actually misleading. The referee is right that three different layers are observed as reported in P5L19-20. The layer mean intensive parameters are given in Table A. The Ångström exponent for the 3 layers maintains a rather stable character with values around 0 suggesting large particles over Potenza. Regarding the lidar ratio, the values decrease with height and confirm the remark that we should not consider a single aerosol type throughout the range.

***Table A:** the mean intensive parameters for the 3 layers observed.*

Layer [km]	*κ_β(355,532)*	*κ_β(532,1064)*	*κ_β(355,1064)*	*κ_α(355,532)*	*S_{355} [sr]*	*S_{532} [sr]*
1.6-2.0	*0.45±0.03*	*0.32±0.03*	*0.37±0.02*	*-0.2±0.2*	*53±8*	*57±8*
2.0-3.5	*-0.02±0.12*	*0.42±0.06*	*0.26±0.04*	*-0.3±0.2*	*48±4*	*53±4*
3.5-5.0	*0.12±0.26*	*0.42±0.18*	*0.31±0.13*	*0.4±0.2*	*46±8*	*41±5*

With respect to the apparent different behavior of extinction and backscatter, we do not think that the increase of the lidar ratio below 2 km is a smoothing effect as both backscatter and extinction coefficient have the same resolution. This increase of the lidar ratio is also followed by a slight increase in the Ångström coefficient (altitudes below 1.6 km) and it might indicate the mixing with local aerosols. It is right mentioning here that below 2 km (see P33-Figure 1) the planetary boundary layer contribution has to be considered. For the tenuous layer above 3.5 km, the lidar ratio has indeed a large slope, however a few hundred meters below 5 km, and with a large statistical error because of the low aerosol concentration.

Besides, we investigated the backward trajectories for the layers of Table A using the HYSPLIT model (Stein et al., 2015). We initiated the model for a 7-day backward analysis and starting height levels the midpoints of the layers and the results are shown in Figure C. The layers 2.0-3.5 km and 3.5-5 km show a similar pattern with the air-masses flying over 2 km and originating source the Saharan desert. The layer 1.6-2.0 km follows a different pathway, the air-masses circulate over the Mediterranean Sea and Algeria and the uptake of marine particles is very likely. This information combined with the information of Table A suggests dust particles for the 2 elevated layers.

However, it is right to mention that this case (and the case in P39-Figure 7) is reported for showing how the classification is done using intensive properties and backward trajectory analyses. All the intensive properties used for setting up the reference dataset and testing dataset are related to elevated layers separated from the planetary boundary layer and local influence.

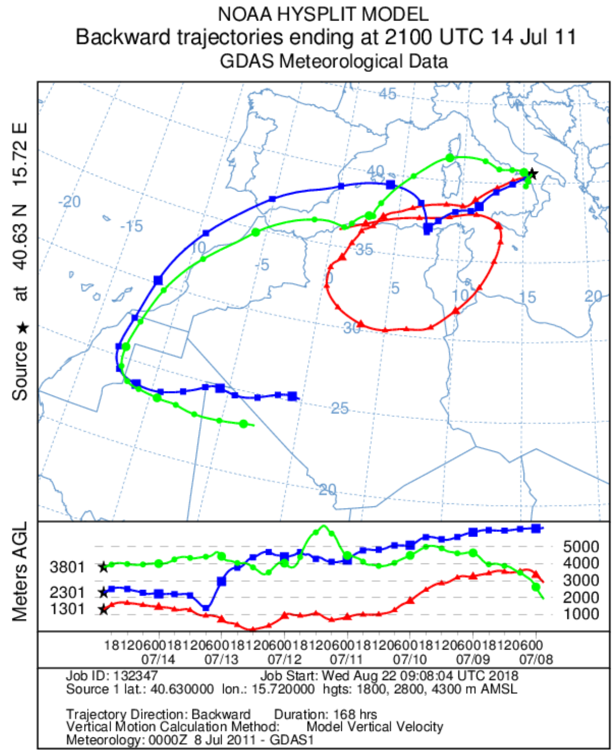

***Figure C:** HYSPLIT backward trajectories for the aerosol layers observed over Potenza, on 14 July 2011, 19:20–22:10 UTC.*

In synthesis, the 3 layers with the synergistic use of intensive properties and backward trajectory analysis allows us to distinguish the different characteristics of the mixed dust layer in the range 1.6-2.0 km and the desert dust higher in the atmosphere (2.0-5.0 km). The intensive properties are measured with good level of uncertainty only in the layer 2.0-3.5 km. In order to avoid confusion, we will focus our comments in the range 2.0-3.5 km where the intensive property analysis is coherent and confirms the existence of a dust layer. The text implemented in P5L22-24 is given below:

The layer 2.0-3.5 km has a constant behavior with the range for the intensive optical profiles indicating the presence of the same type of particles. The mean values of all optical parameters in the range are calculated: lidar ratios of 48±4 sr at 355 nm and 53±4 sr at 532 nm and Ångström exponents (κ_β(355,1064), κ_β(532,1064), κ_β(355,532), κ_α(355,532)) of -0.3–0.4 were found.

Furthermore, the sentence P13L31 is changed to:

For the case in Sect. 2.1, the automatic algorithm labelled the aerosol layer as dust, D_M=1.2 and the normalized probability 55 %.

The same kind of discussion applies for P39-Figure 7, where the vertical raw resolution is 7.5 m and the effective resolution of the extinction coefficient profiles is given in Figure D. The resolution for both wavelengths degrades with height (faster for 532 nm). For the height range 1.2-1.9 km, the highest resolution is 240 m for 355 nm while for 532 nm varies from 240 m to 480 m. Higher in the atmosphere, the resolution degrades with height (faster for 532 nm) and becomes constant at 4.7 km for 355 nm (3.8 km for 532 nm).

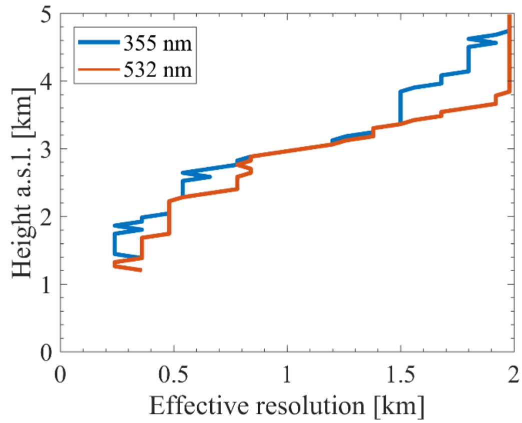

Figure D: *The effective resolution of the extinction coefficient at 355 nm (blue line) and 532 nm (red line) for the Athens lidar on 22/05/2014, 20:28–21:28 UTC.*

Similar to Figure B, Figure E shows the same panels as for P39-Figure 7 along with the smoothed backscatter profiles (panel b) that were used for the calculation of the lidar ratio. The next sentence is implemented in the revised manuscript:

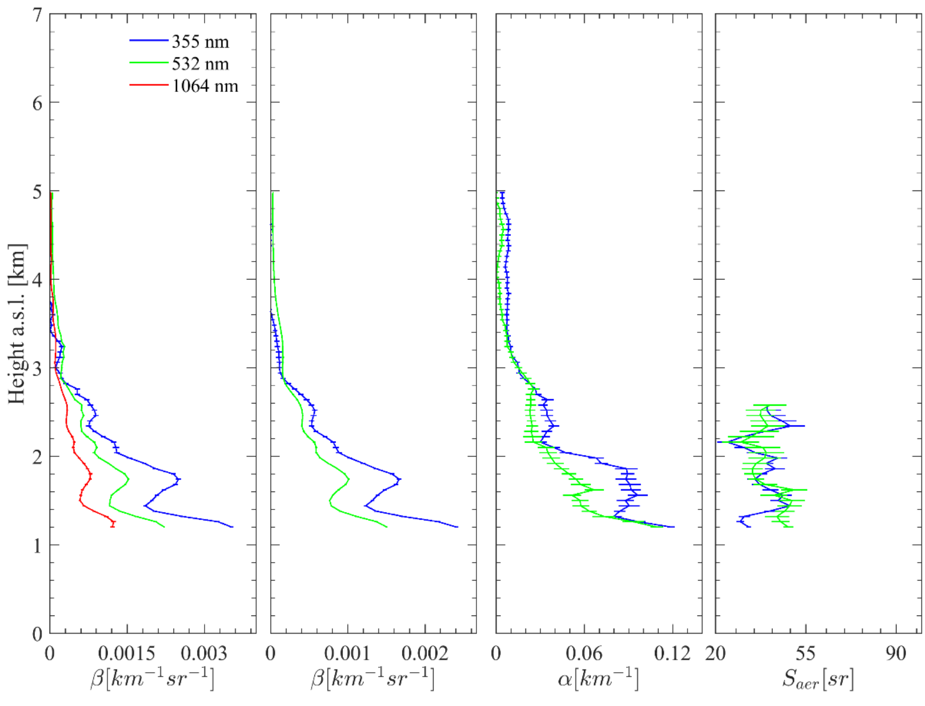

Figure E: *Optical profiles measured in Athens, on 22 May 2014, 20:28–21:28 UTC with a multiwavelength Raman lidar. From left to right, (a) the particle backscatter coefficient with the full resolution, (b) the smoothed particle backscatter coefficient, (c) the particle extinction coefficient, and (d) the particle lidar ratio. The error bars correspond to the standard deviation.*

The effective resolution of the extinction coefficient profiles varied from 240 m to 780 m using the method described in Pappalardo et al. (2004).

2. Would you say that there is a possibility for error in the determination of "truth" aerosol types? If so, it would be good to see some discussion of that.

Aerosol classification as described in Section 2.1 is based on a descriptive analysis of the retrieved optical properties and the model simulations, and it is a qualitative method of type assignment. Thus, there is inherent possibility of error in the determination of the true aerosol type. This error, if made, propagates into the automatic algorithm and the predicted aerosol class might deviate from the "truth" aerosol class.

As discussed in Sections 2.1 and 3.2, the aerosol type was decided manually by using a set of different tools (transport and trajectory models, as well as other observational data). Thus, optical properties for different aerosol types observable over Europe were derived. For the selected aerosol types, P27-Table 2 highlights the characteristics of each type and the values concur with several aerosol typing studies, such as Müller et al.(2007a), Burton et al. (2012, 2013, 2014), Gross et al. (2015), and Giannakaki et al. (2016).

Nevertheless, there is the possibility for error in the aerosol type definition. In supervised learning techniques, the aerosol classes must be well defined, differently the predictions of the predictive algorithm will suffer from the wrong definition. For example, this aspect can be observed in the definition of the Smoke and Volcanic category. The former refers to biomass burning particles with small size and very high lidar ratios whereas aged smoke plumes are expected to have different characteristics. The later refers only to fresh volcanic emissions and the optical characteristics are similar to the desert dust particles. Moreover, the dust type considers plumes from the Saharan desert and not the Arabian where the lidar ratio tends to be smaller. All these will lead the algorithm to misclassify the aerosol types not taken into account and allocate them to a wrong type. However, the flexibility of the algorithm allows us to circumvent it by inserting more and/or better defined aerosol types. The next lines are inserted in Section 3.2.1:

However, aerosol classification is based on an interpretative analysis of the retrieved optical properties and the model simulations and, it is a qualitative method of type assignment. Thus, there is inherent possibility of error in the determination of the true aerosol type. This error, if made, propagates into the automatic algorithm and the predicted aerosol class might deviate from the "truth" aerosol class.

3. I also have a particular question about the interpretation of the influence of marine aerosol in the two case studies that are discussed at length. FLEXPART, Figures 3 and 8, seems like a very nice tool for information on aerosol source. Figure 3 seems to show a large fraction of the incidence below 2 km (a lot of the green and yellow) as being in the Mediterranean Sea, but in the discussion, no mention is made of marine influence and the case is described as "pure dust". On the other hand, in Figure 8, an apparently lesser proportion of the trajectories below 2 km are seen over the Black Sea and the Caspian Sea, and this layer is described as a mixture "enriched with marine particles during their overpass over the Black Sea". Can you clarify how we can know that there is marine influence in one case and not the other? It seems that it would be particularly difficult to say definitively when aerosol types are "pure" rather than mixed, using this method. Can you clarify whether there are additional factors that go into these judgements besides the FLEXPART tool and what uncertainties are associated with those judgements?

The manual typing as described in Section 2.1 is not a simple issue and, of course, leaves room to questionable type assignment. The backward trajectory analysis is used synergistically with the lidar optical properties and the model outputs suffer from the high error on the path (increasing with the path itself) and the source term *assignment. Therefore, the model simulations have to be used together with the observed lidar optical properties for providing a reasonable aerosol typing.*

The FLEXPART plot indicates the possibility of marine influence in the identified layer, however the lidar ratio values being over 50 sr indicates low or no presence of maritime aerosol particles. Furthermore, for this specific case, particle linear depolarization ratio measurements are available, however these measurements are not included in the database and therefore not reported in the submitted manuscript. Figure F shows the particle linear depolarization ratio where the values are over 0.2 and, thus, confirming our hypothesis of aspherical particles.

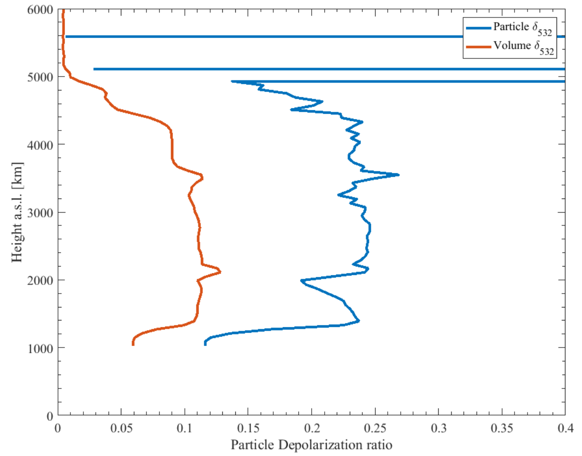

***Figure F:** Particle (blue line) and Volume (red line) linear depolarization ratio at 532 nm for the MUSA system in Potenza on 14/07/2011, 19:20–22:10 UTC.*

In order to avoid confusion, we rephrased the sentence in P5L30-31 and moved P5L24-25 to end of it:

The dust-prone area of northern Africa (Morocco and northern Algeria) along with the Mediterranean Sea are most likely the sources of the observed layer and suggest a mixture of dust and marine particles. The combined information of the backward trajectory analysis and the intensive properties values indicate the presence of dust particles and they are in accordance with the typical dust values observed over Potenza (Mona et al., 2014).

P40-Figure 8 shows a complicated scene with multiple aerosol sources with low probability of marine influence. Unfortunately, there was a misinterpretation of the FLXEPART plot. The sentence in P14L12-15 has been rephrased and implemented in the text:

Therefore, the path of the air masses arriving over Athens suggests a mixture of dust and biomass burning particles, originating from the arid areas of the Aral Sea, as well as the agricultural fires in former Soviet Union countries (Papayannis et al., 2016).

4. <u>Section 3.2.4. When you add particle depolarization as a classification variable, I think you should still keep the original three variables. You have already shown that all three variables are sufficiently independent to be useful, so adding an independent fourth variable would be expected to produce the best classification method. I'm not following why you avoid using more than 3 variables.</u>

The reason that we removed a classifying parameter was that we wanted to compare in a fair and informative manner the two setups (i.e., the 3+2 and the 3+2+1). Regarding this remark, the text and figures have been corrected in the revised version of the manuscript.

The main features of the Figure G (same as P38-Figure 6) remain almost the same. The difference between with and without depolarization ratio of the error rate of LOOCV is bigger, thus, indicating a better performance when the 4 classifying parameters are used.

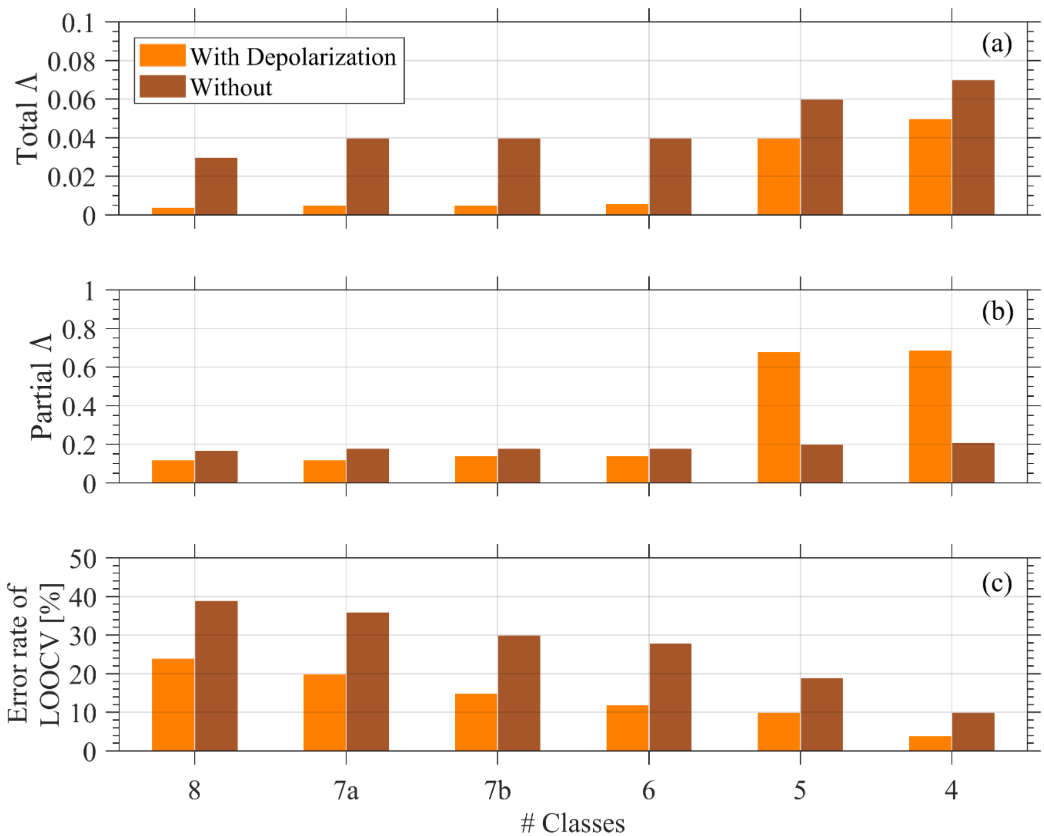

Figure G*: Bar plots showing a) the total Λ, b) the partial Λ, and c)error rate of LOOCV when comparing the training of the algorithm with (i.e., S_{532}/S_{355}, S_{532}, and $\kappa_\beta(355,1064)$) and without (i.e., $\delta_{aer,532}$, $S_{aer,532}$, and κ_β) particle linear depolarization values. For the partial Λ, the brown bars correspond to the backscatter-related Ångström exponent and orange one to the particle linear depolarization ratio because they represent the most significant classifying parameter in the classification.*

Figure H (same as P41-Figure 9) shows a different behavior when compared to the graph reported in the first version of the manuscript. The prediction accuracy when depolarization is used remains almost stable (79%-85%) for all considered classes and do not increase almost monotonically when aerosol types are merged. On the contrary, for 4 classes the prediction accuracy is the lower among the other classes, however, still it is very high. The graphs and the corresponding discussion is changed accordingly.

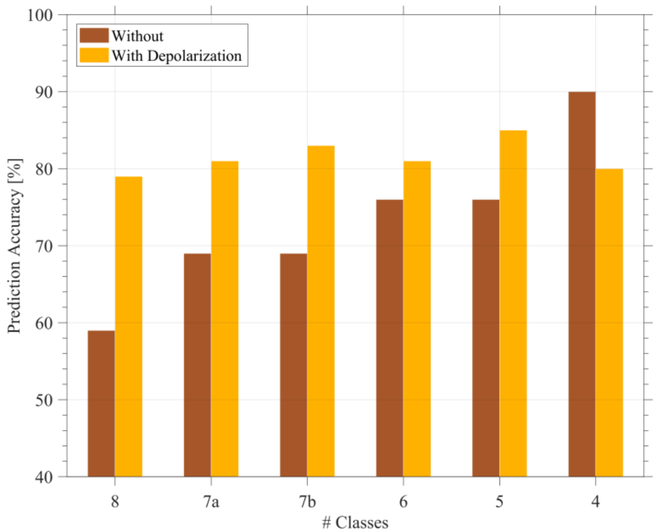

Figure H: Prediction accuracy for the different aerosol classes and with/without depolarization information.

The next lines have been inserted in the revised version of the manuscript:

P13L6-11: **In this case, the particle linear depolarization ratio was added to the classifying parameters. Values within the aerosol type range were randomly assigned to each sample and the Λ distribution was calculated. Total Λ is 0.004. The value of partial Λ for κ_β, S_{532}, S_{532}/S_{355} and δ_{532} are 0.55, 0.34, 0.52 and 0.12 respectively. The values found for the partial Λ confirm the δ_{532} as the most important classifier for the considered dataset. For the rest aerosol groups the total and partial (for depolarization ratio) Λ are 0.005 and 0.14 respectively (7[a] classes), 0.005 and 0.12 (7[b] classes), 0.006 and 0.14 (6 classes), 0.040 and 0.68 (5 classes), and 0.050 and 0.69 (4 classes).**

P15L12-15: **With depolarization ratio information, the accuracy for 8 classes equals to 79 % and exceeds the 80 % for the rest aerosol classes. When comparing the accuracy of the model with and without depolarization ratio, it appears to be significantly higher until 6 classes where, further, the discrepancy diminishes (<10 %) and becomes smaller for 4 classes.**

P16L19-21: **Besides, the training of the algorithm with literature depolarization ratio values decreased the error rate of the leave-one-out cross validation from 24% (8 classes) to 4% (4 classes). Furthermore, the predictive accuracy increased and remained for all the aerosol classes around 80% (for 8 classes: 79%, for 7a: 81 %, for 7b: 83 %, for 6: 81 %, for 5: 85 %, and for 4: 80 %).**

5. <u>P4 L14 The idea that the spectral ratio of lidar ratio indicates smoke aging is still a hypothesis, based empirically on a small number of suggestive cases. Describing this relationship as "robust" overstates the case, I think. Not all of the references actually support the statement. For example, Samaras et al. 2015 have no measurements related to smoke age, but rather take it as given, substituting the spectral ratio of lidar ratios as a proxy for smoke age. Please don't use a reference to support a hypothesis that merely made use of the hypothesis (at least not without more explanation). Anyway, the current manuscript</u>

doesn't relate to smoke aging. You could easily remove the statement and avoid controversy. At least don't overstate it and please remove references that do not really support the statement.

The sentence has been rephrased and the non-relevant references are removed. The next sentence substitutes the phrase in P4L13-15:

This quantity has shown the ability to characterize the ageing status of smoke particles as well as the spectral dependence of aerosol (Müller et al., 2007a; Alados-Arboledas et al., 2011; Nepomuceno Pereira et al., 2014; Nicolae et al., 2013).

6. Should the "clean" in the label "clean continental" be taken literally? The description of the clean continental category is obliquely defined here as a mixture of polluted continental and clean marine aerosol and indeed the data in Figure 5 also seem to support its interpretation as a mixture of the two. I think this is an interesting way to think about this type, perhaps much more useful than the standard way of thinking of it as a type that, unlike all the others, is defined by an extensive aerosol property (low aerosol loading). Any comment on this? Would a case that has the intensive properties of the clean continental class but a significant amount of aerosol optical depth (so therefore not particularly "clean") be considered "clean continental" in your analysis?

The clean continental category, as defined in P6L20-23, presents the mixture of anthropogenic aerosols with natural sources. It is true that the clean continental subtype occupies the space between polluted continental and mixed marine (P34-Figure 5), however this aerosol type is observed in continental EARLINET sites far from maritime aerosol sources (e.g., Bucharest, Leipzig). The manual aerosol typing made by Schwarz (2016), part of the reference dataset, assigns an aerosol layer as clean continental when the aerosol concentration is low by means of optical depth. Based on this definition, the label "clean" can be taken literally. However, the automatic typing procedure does not take into account any extensive parameter, hence not "clean" aerosol layers might be classified as clean continental.

7. P9-10. The information about Mahalanobis distance is basically repeated from earlier work. You could simplify by referencing Burton et al. 2012 and noting the different thresholds.

The section 3.1 has been reshaped accordingly. The section reads below:

We developed an automated typing method, based on the work of Burton et al. (2012), but modified it in order to be compatible with the database of EARLINET. Two major steps are identified in the method proposed: the training (Sect. 3.2), and the testing (Sect. 4.1) phase. The first step consists of the following procedures. As described in Sect. 3.2.1, well characterized aerosol samples are manually separated into classes based on their physical characteristics; the set of classes constitutes the reference dataset. This procedure involves the determination of each observed aerosol layer location and the estimation of mean layer intensive optical properties. Based on this analysis, the classifying parameters that provide the required information for a better discrimination of the aerosol type are selected (Sect. 3.2.2). Next, in order to estimate how accurately a predictive model will perform, the reference dataset is split into training and validation datasets, and the application of the classifier is evaluated (Sect. 3.2.3). Sect. 3.2.4 describes the inference of characteristic depolarization values in the algorithm with the intention to increase the prediction of the model. For the second step, already pre-classified EARLINET data are used to assess the performance of the automatic typing procedure. Figure 1 illustrates the sequence of the proposed methodology starting from the setting of the training dataset, up to the assessment of the learning success during the testing phase.

The Mahalanobis distance of an observation from an aerosol class is estimated, and is assigned to the aerosol class for which the distance is minimum. Two screening criteria are applied to the minimum distance following the procedure of Burton et al. (2012). The methodology uses 3 and 4 classifying parameters and the minimum accepted distance for a measurement to be labelled is 4 and 4.3 respectively. Moreover, the normalized probability of the aerosol class needs to be higher than 50 %.

8. <u>P10, L7. Similar probabilities for two different classes do not indicate mixing between those two classes. This only reflects that those two classes are close to each other in your measurement space. For example, any point that is close to your "smoke" class is also close to your "polluted continental" class because the classes are close to each other. If it actually is smoke plus a little bit of marine influence, the 2nd closest class will still be "polluted continental", not marine.</u>

Thanks for this comment. This sentence is removed.

9. <u>Section 3.2.2 classifying variables selection. I think choosing variables solely based on Wilks' partial lambda may not catch everything. It may be that different variables have more power to separate different subsets of classes. For example, depolarization obviously has a lot of power to separate dust classes and almost no power to separate non-dust classes. If you had a variable that helped separate smoke from polluted continental, even partially, but did nothing else, it may have a poor partial lambda but it would nevertheless be extremely valuable. I suggest a plot similar to figure 10 in Burton et al. 2012 as a way to understand more thoroughly what each variable contributes to separating the classes. By looking at the variability of each variable within each class you can see where the overlaps are in every dimension. Your figure 5 also does this but it's usefulness maxes out at 2 dimensions, since it is very difficult to visualize more than 2 dimensions in a literal space. Your Wilks' lambda analysis suggests that the ratio of lidar ratios has significant discriminatory power, but Figure 5 does not reveal how (that is, whether some sets of classes that look like they overlap might be distinguished by the 355 nm lidar ratio or the spectral lidar ratio). It would be nice to have a visualization that answers that question. This would also address the question of whether spectral lidar ratio separates smoke and pollution aerosol (Muller et al. 2007a) which would be an interesting discussion in itself.</u>

The starting point of this work was a 3+2 Raman lidar configuration. Therefore, we sought for intensive parameters that are type sensitive and can be used for an effective aerosol characterization. The intensive parameters were presented in Section 2 and type specific values were given in Section 2.2. The Wilks' lambda analysis performed in Section 3.2.2 was made on the basis of the available wavelengths and the already pre-selected classifying parameters: the lidar ratio, the Ångström exponent, and the spectral ratio of the lidar ratio.

The bar chart in Figure I show the median (black dot), 25-75 percentile (box), and 5-95 percentile (whiskers) for each of the aerosol types and each classifying parameters. The discriminatory power of the Ångström exponent and lidar ratio is apparent whereas the ratio of the lidar ratios performs the worst among the three, in agreement with the Wilks' lambda analysis. The polluted and mixed dust yield the highest variability in the ratio of the lidar ratios that spans a wide range with median values slightly over 1. The rest of the categories have median values below 1 with dust and volcanic being the types with values almost 1. This behaviour can be seen as a confirmation of their spectral independence. Regarding Smoke and Polluted Continental there seems to be no evidence that this variable can distinguish them. However, this finding is a further confirmation of the comment #2.

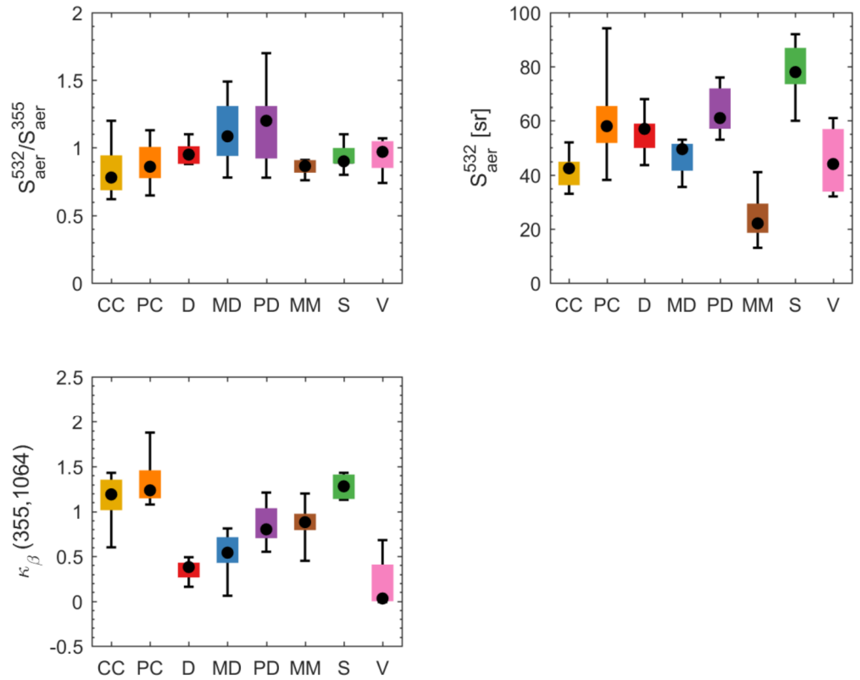

Figure I*: Bar plots and whiskers show the median (horizontal line), 25-75 percentile (box) and 5-95 percentile (whisker) of the three classifying parameters: κ_β(355,1064), S_{532}, and S_{532}/S_{355}.*

10. <u>P16, Perhaps you could discuss more explicitly the tradeoff between more classes and less classes. All of your statistics (except Wilks' lambda) seem to show better performance with fewer classes, but that could be taken to an extreme. That is, with only one class, there would be no errors at all! How do you address this tradeoff?</u>

Typing in multiple classes and typing accuracy are two conflicting aspects. The number of classes as well as the typing accuracy depends on the specific needs. Currently typing methodologies provide different typing outputs with different level of accuracy: e.g., MISR (Multi-angle Imaging SpectroRadiometer) algorithm provides aerosol classification into 5 classes with high accuracy but also a finer typing (74 classes) with degraded accuracy (Kahn et al., 2015).

This could be an approach leaving to the specific user the possibility to select appropriate balance for his own application. Another possibility is to find a compromise between degrading accuracy and gaining insight into the aerosol type. In this direction, we indicated into the conclusions that for a 3β+2α setup, 6 aerosol classes can be used for the aerosol typing whilst for a 3β+2α+1δ setup a finer classification – with 7 aerosol classes – can be made. Whenever better performance in terms of accuracy is requested the number of aerosol classes can be reduced to 4 because of the similarities of the physical characteristics of the aerosol classes. The choice of 4 classes is also a good analogy for the proposed aerosol typing by Baars et al. (2017).

The next phrase is inserted in Section 3.2.3:

It should be mentioned that the typing in multiple classes and typing accuracy are two conflicting aspects. The choice of 8 aerosol classes appears to be sufficient to describe the major aerosol components, however ostentatious for a 3+2 lidar configuration. 4 classes, on the other hand, provide a coarse aerosol characterization and the prediction accuracy of the algorithm is expected to increase.

11. <u>Do you plan to share these aerosol typing results publically? What about the training database of manually typed samples? Also please include links to the EARLINET database.</u>

The data used into this paper are freely and publicly available on the European Research Infrastructure for the observation of Aerosol, Clouds, and Trace gases (ACTRIS) website (https://www.actris.eu/). Moreover, the complete EARLINET dataset for 2000-2015 period is currently on the way for publication on the CERA database. This information will be added, if available, before the final publication of the paper.

Additionally, the reference and training datasets will be publicly available as well: a special dataset will be available through EARLINET/ACTRIS web page. In addition, this dataset will be part of a larger initiative about aerosol typing within AEROSAT (International Satellite Aerosol Science Network; http://www.aero-sat.org/), in which it is planned to set up an extensive reference dataset for aerosol typing investigation.

Typos and requests for clarification:

1. <u>P2, last sentence: Both cluster analysis techniques and supervised classification techniques need the number of groups as input.</u>

Yes, it is true that for cluster analysis it is needed the number of groups as input, however the identities of groups are not known in advance in contrast with the classification analysis (see Wilks 2006, page 529). The P2L32-33 sentence is rephrased:

Whereas in cluster analysis, the groups are not known beforehand and the classifier is tasked with it.

2. <u>P4 L7: "operates" should be "operate"</u>

Ok, fixed.

3. <u>P5, L18: talks about the region of incomplete overlap. What altitude does this go up to?</u>

The full overlap height for the 1064 nm channel is approximately 405 a.g.l. m (Madonna et al., 2015). This information is implemented in the text:

MUSA has a full overlap at around 1,15 km a.s.l. for 1064 nm (Madonna et al., 2015).

4. <u>P5, L22-23: the sentence "the aforementioned layers" is unclear and should be reworded. I think you are suggesting that the intensive properties are approximately constant throughout each layer, but that does not really appear to be true, so perhaps I'm misunderstanding the wording.</u>

Based on the comment #1, the sentences P5L22-24 now read:

The layer 2.0-3.5 km has a constant behavior with the range for the intensive optical profiles indicating the presence of the same type of particles. The mean values of all optical parameters in the range are calculated: lidar ratios of 48±4 sr at 355 nm and 53±4 sr at 532 nm and Ångström exponents (κ_β(355,1064), κ_β(532,1064), κ_β(355,532), κ_α(355,532)) of -0.3–0.4 were found.

5. P5, L24 and throughout: there are four Angstrom exponents discussed but often the text refers to "Angstrom exponent" as if there is only one. Please clarify which one you mean. Likewise, it should be specified which wavelength is meant when "lidar ratio" is used.

For P5L24, we use "Ångström exponents" for all the available 4 sets appearing in Figure 2 – κ_β(355,1064), κ_β(532,1064), κ_β(355,532), κ_α(355,532). The Ångström exponent and lidar ratio wavelengths, whenever it was needed, are now reported. The following sentences are implemented in the text.

P8L23: **lidar ratios at 532 nm of 50–65 sr**
P8L27-28: **Pappalardo et al. (2004) and Wang et al. (2008) reported lidar ratios of 50–60 sr at 355 nm and backscatter-related Ångström exponent (355, 532) of 2.4.**
P5L24: **Ångström exponents (κ_β(355,1064), κ_β(532,1064), κ_β(355,532), κ_α(355,532)) of 0–0.4 were found.**
P10L28: **Ångström exponent for the available.**
P10L30: **The Ångström exponents (κ_β(355,1064), κ_β(532,1064), κ_β(355,532), κ_α(355,532)) for mixed dust are.**
P10L31: **Ångström exponents lie.**
P10L33: **with mean Ångström exponent from all the available variables around ∼1.4 and ∼1.3 respectively.**
P11L1: **81±16 sr and 78±11 sr for 355 nm and 532 nm respectively.**
P11L3: **Ångström exponents (κ_β(355,1064), κ_β(532,1064), κ_β(355,532), κ_α(355,532)) are.**
P11L5: **Ångström exponents (κ_β(355,1064), κ_β(532,1064), κ_β(355,532), κ_α(355,532)) in the range 0.8–1.0**
P12L3: **S_{532} and κ_β(355,1064)**
P13L16: **lidar ratio at 532 nm.**
P14L5: **Within this layer the mean value of backscatter-related Ångström (355,1064) exponent is 0.9±0.1**
P38-Figure 6: **κ_β(355,1064).**

6. P6 L23. I recommend taking more care about using the word "absorbing". It seems that "more absorbing" is here used as a synonym for "higher lidar ratio", but lidar ratio depends on particle size and other factors as well as light absorption and is really not as direct an indicator as this language suggests. Also P11 L4 and L14 (continental pollution is not necessarily absorbing); P8 L5 (smoke is often absorbing but not always highly absorbing especially when aged); and perhaps elsewhere.

It is true that the lidar ratio depends on many factors and it is not a direct indicator of light absorption. The following sentences are changed and implemented in the manuscript.

P6L23: **The clean continental, therefore, differentiates from the polluted continental type due to lower lidar ratio values.**

P11L4: **This characteristic separates clean continental from polluted continental as the particles yield lower lidar ratio values.**

P11L14: **Two pathways were followed, first, the smoke and the polluted continental categories were grouped into the more generic type of small with high lidar ratio values.**

P8L5: **Generally, smoke particles are relatively small and spherical that produce low depolarization, high Ångström exponents, and large lidar ratios (Amiridis et al., 2009; Baars et al., 2012; Giannakaki et al., 2016).**

7. P7, L14. Delete "mainly". Although the variability in the lidar ratio is a "hot topic" there is also significant variability in, for instance, extinction Angstrom exponent.

Deleted.

8. P7, L31. The suggestion for CALIPSO to add a dust+marine type was made several times before 2016 also, for example Kim et al. 2013, Burton et al. 2013, Rogers et al. 2014.

Thank you for this suggestion. These literature references are implemented in the document and the sentence reads as follows:

Several studies (Burton et al., 2013; Kim et al., 2013; Rogers et al., 2014; Papagiannopoulos et al., 2016a) have indicated that this mixture is important and suggested its inclusion in the CALIPSO retrieval scheme for improving the accuracy of aerosol backscatter and extinction coefficient profiles.

9. P8 L10 and in the references: I think Pereira should be Nepomucino Pereira (that is, the first author appears to have a two-part surname).

Yes, it is fixed. The reference becomes:

Nepomuceno Pereira, S., Preißler, J., Guerrero-Rascado, J. L., Silva, A. M., and Wagner, F.: Forest Fire Smoke Layers Observed in the Free Troposphere over Portugal with a Multiwavelength Raman Lidar: Optical and Microphysical Properties, Scientific World Journal, 2014, doi:10.1155/2014/42183, 2014.

10. P9 first paragraph and third paragraph: There are a few places, including these 2 paragraphs, where the wording is awkward with several errors in English language usage that make them hard to understand. Please reword for clarity

Following specific comment #7, the first paragraph is reworded and merged with the second paragraph as follows:

We developed an automated typing method, based on the work of Burton et al. (2012), but modified it in order to be compatible with the database of EARLINET. Two major steps are identified in the method proposed: the training (Sect. 3.2), and the testing (Sect. 4.1) phase. The first step consists of the following procedures. As described in Sect. 3.2.1, well characterized aerosol samples are manually separated into classes based on their physical characteristics; the set of classes constitutes the reference dataset. This procedure involves the determination of each observed aerosol layer location and the estimation of mean layer intensive optical properties. Based on this analysis, the classifying parameters that provide the required information for a better discrimination of the aerosol type are selected (Sect. 3.2.2). Next, in order to estimate how accurately a predictive model will perform, the reference dataset is split into training and validation datasets, and the application of the classifier is evaluated (Sect. 3.2.3). Sect. 3.2.4 describes the inference of characteristic depolarization values in the algorithm with the intention to increase the prediction of the model. For the second step, already pre-classified EARLINET data are used to assess the performance of the automatic typing procedure. Figure 1 illustrates the sequence of the proposed methodology starting from the setting of the training dataset, up to the assessment of the learning success during the testing phase.

The third paragraph is changed:

For the second step, unclassified EARLINET data (the testing dataset) is categorized using the reference dataset. Besides, the testing dataset has been classified following the method shown in Sect. 2.1 and, hence, compared against the output of the automatic procedure. Figure 4 illustrates the sequence of the proposed methodology, starting from the setting of the training dataset up to the assessment of the learning success during the testing phase.

11. <u>P12 L28, I think you mean Burton et al. 2015, not 2014. Burton, S. P., Hair, J. W., Kahnert, M., Ferrare, R. A., Hostetler, C. A., Cook, A. L., Harper, D. B., Berkoff, T. A., Seaman, S. T., Collins, J. E., Fenn, M. A., and Rogers, R. R.: Observations of the spectral dependence of linear particle depolarization ratio of aerosols using NASA Langley airborne High Spectral Resolution Lidar, Atmos. Chem. Phys., 15, 13453- 13473, 10.5194/acp-15-13453-2015, 2015.</u>

Thank you. Fixed.

12. <u>P13, L25 & L27, and throughout. When you count measurements or samples, what defines a single measurement, given that the lidar systems operate basically continuously? Please discuss how you select data and discuss what criteria are used in data selection and how much averaging is done.</u>

EARLINET is a research network born in 2000 on the basis of bringing together the existing research lidars around Europe. At that time, none of the members could provide continuous measurements. Nowadays, there are some systems working 24/7 unattended, but EARLINET is working on a measurements schedule (see P4L1). Furthermore, measurements are performed during CALIPSO site overpasses and special events, such as biomass burning episodes, dust advection, and volcanic eruptions. To accommodate these kind of events, it was agreed to have some measurements lasting 2-3 hours in order to monitor the evolution of each specific event. Then, it is responsibility of the stations to provide aerosol profiles better representing the observed atmospheric aerosol conditions: i.e., deciding the time window and the time averaging needed for the lidar signal inversion.

Throughout the paper, we refer to measurements as a set of aerosol optical profiles reported in the EARLINET database and corresponding to the same temporal window, which typically spans about 1 hour of temporal integration.

For making clearer this point we added the following sentence in P4L6:

Throughout the paper, we refer to measurements as a set of aerosol optical profiles reported in the EARLINET database that correspond to the same temporal window, which typically extends for about 1 hour.

We refer instead to samples as identified layers. We changed samples into layers in the related text.

13. <u>P13, L25 & L27. Also, the numbers in this paragraph are confusing. If there are only 42 measurements available with three wavelengths, how are there 47 samples? Of if you don't need all 3 wavelengths, then why only 47 instead of 157?</u>

For the considered period there are 157 measurements (see definition above). 42 of which provide 3 (at 355 nm, 532 nm, and 1064 nm) backscatter coefficient profiles and 2 (355 nm and 532 nm) extinction coefficient profiles (this reduced number is due to low SNR for some optical properties especially extinction with inhibited the retrieval but most to the fact that in daytime conditions Raman technique does not allow the extinction retrieval at all). Out of the 42 measurements, 47 aerosol layers have been identified in the free troposphere and a reasonable typing was possible.

14. <u>P14-15, discussion of recall, precision and accuracy (also error rate from the earlier discussion). The description of these variables could be made clearer and it should be specified that recall and precision refer to specific classes. I think the use of the terms false negative and false positive contributes to confusion rather than clarity; the descriptions are too similar to distinguish them. I think you are saying that recall for a particular class indicates the number of correct identifications divided by the number of actual instances of that class. Precision for a particular class indicates the number of correct identifications of that class divided by the number of times when that class was predicted, whether rightly or wrongly.</u>

Am I understanding it right? And what is "accuracy"? Is it defined the same as "error rate" or does "accuracy" only count the instances that are classified and not the ones that are unclassified due to overlap between classes? Are all four of these statistics independently useful, or could it be simplified by using fewer?

Am I understanding it right? *Yes, this is what recall and precision are.*

And what is "accuracy"? Is it defined the same as "error rate" or does "accuracy" only count the instances that are classified and not the ones that are unclassified due to overlap between classes? Accuracy is defined as 1-Error Rate (as introduced in Section 3.2.3), and it is demonstrating how often the classifier is correct.

Are all four of these statistics independently useful, or could it be simplified by using fewer? *In order to reduce the number of the statistics, the accuracy is replaced by the error rate throughout the document.*

The description of the error rate in P12L20-21 is changed and implemented in the text:

The error rate is estimated as a percentage of all incorrect predictions divided by the total number of the reference dataset, and is equivalent to 1 minus accurary. Values near zero show high predictive performance while values near one show low predictive performance.

The recall and precision relate to one class independent of any other classes. Error rate or accuracy is not a reliable metric for the assessment of the real performance of the classifier, because it may produce misleading results. That is, the number of observations vary greatly and this effect can be seen in P30-Table 5. We decided to keep the metrics recall and precision, however we provide a clearer description and the formulas removed. The P14L25-30 & P15L1-8 is restructured and the next sentences are introduced in the document:

… is already known. Recall of an aerosol group is defined as the number of correctly predicted cases over the number of correctly plus the number of incorrectly predicted cases. Recall can be thought as the model's ability to predict the specific aerosol class. Precision of an aerosol class is defined as the number of correctly predicted cases over the number of correctly predicted cases plus the number of incorrectly predicted cases that belong to this aerosol class. In other words, given the prediction of a specific class, what is the probability of being correct?

15. P15 L9 "cannot be evaluated". It seems that even though they don't appear in the training set, they were occasionally predicted by the method. If so, perhaps reporting ow often that happened would be useful.

This comment refers to P15L30. The frequency of detection for MD (PD) is 18% (4%) for 8 classes, 15% (3%) for 7^b classes, 17% (3%) for 7^a classes, and 15% (3%) when 3 classifying parameters are used. The algorithm predicted as MD only dust cases with mean S_{532}=47±2 sr, S_{532}/S_{355}=1.1±0.1 and κ_β(355,1064)=0.4±0.1; thus the lower lidar ratio and the over 1 value of the ratio of the lidar ratio assigns these cases as dust. The PD case refers to a PC case with S_{532}=69 sr, S_{532}/S_{355}=1.2 and κ_β(355,1064)=0.9.

The frequency of detection for MD is 0% for all classes when particle depolarization ratio is added. The frequency for PD is 17% for 8 classes, 13% for 7^b classes, 17% for 7^a classes, and 13% for 6 classes. The PD refers to a case with low particle depolarization ratio. Following this remark, we added the next sentences in the revised document:

Note that mixed dust and polluted dust aerosol types are not reported in the tables due to the fact that they are not present in Table 5 and these parameters cannot be evaluated. The frequency of detection for MD (PD) is 18% (4%) for 8 classes, 15% (3%) for 7^b classes, 17% (3%) for 7^a classes, and 15% (3%) when 3 classifying

parameters are used. The algorithm predicted as MD only dust cases with S_{532} around 45 sr and S_{532}/S_{355} over 1. The PD case refers to a PC case with κ_{β}(355,1064) lower than 1. The frequency of detection for MD is 0% for all classes when depolarization ratio is added. The frequency for PD is 17% for 8 classes, 13% for 7^b classes, 17% for 7^a classes, and 13% for 6 classes. The wrongly classified cases have depolarization ratio around 20%.

16. Table 1. I don't understand what the value in parentheses represents.

The number in parentheses indicate the mixtures of more than two aerosol types. These numbers will be removed because they do not add to our discussion.

17. Table 2. Am I interpreting this correctly, that smoke is never predicted by the classification methodology? Any comments about implications of that?

We think the comment refers to P31-Table 6. The smoke cases examined show optical properties similar to polluted continental, hence the Mahalanobis distance classifier labels them as polluted continental. The mean lidar ratio at 532 nm is 61±5, the β-related Ångström coefficient (355, 1064 nm) is 1.5±0.2, and the spectral ratio of the lidar ratios is 1.15±0.08.

We think that the high values of lidar ratio considered in the reference dataset do not allow the correct classification of smoke layers and as a consequence these smoke layers are typed as polluted continental. We acknowledged in the manuscript that well-defined aerosol types are paramount for the correct classification and, for instance, it is needed to split the smoke category into fresh and aged biomass burning.

18. Table 2. I think PC+C in the 7th column on the top should be PC + S, and I think the 2nd CC in the 4th column on the bottom should be S.

This was a typo and it is fixed.

19. Figure 4. What does "trained classifier" in the middle of the flowchart mean? How is it different from "typing procedure"?

By trained classifier we consider the whole iterative process whereby we build the best possible classifier, i.e., the Section 3. Whereas the typing procedure refers to the labeling of the instances using the Mahalanobis distance classifier.

We consider changing the word "classifier" with "classifying parameter" in the P13L6 and P13L9 as it might confuse the reader.

Referee Comment #2

1. <u>Section 2.2.1 to 2.2.5: You show an extensive overview of the intensive optical properties for the different aerosol types. However, it would be valuable if you could provide further information if the different intensive properties have been derived from the same studies. For using those in the classification the influence of miss-classifications and mixtures should be minimized and the measurements providing a multitude of intensive optical properties should have a larger weight. Additionally you overview mainly focuses on the information presented by Burton et al. or from EARLINET measurements. It would be also valuable to include further measurements (e.g. closer to source regions or after long-range transport) to better differentiate between possible influences of transport or mixture.</u>

The proposed methodology aims to be an EARLINET stand-alone typing methodology. EARLINET collected observations, since 2000, provide an insight of the aerosol types occurring over Europe. Therefore, the methodology is set up and the aerosol classes defined. The aerosol properties are shown to vary with location and aerosol type. For example, as reported also in the paper, desert dust can have different properties depending on the source region (Arabian versus Saharan dust particles). This means that the automatic typing can be more efficient if set up at regional/continental level.

On this basis, Section 2.2 gives an overview of the characteristics of each aerosol type along with intensive optical properties from various literature references. The goal of this section is to introduce the aerosol classes that the automatic typing is based upon and to provide typical values of the intensive parameters. These intensive parameters correspond to the selected classifying parameters and act as a reference to our training dataset. Besides, the majority of the literature references come from EARLINET observations and reflect the variability of the aerosol properties over Europe. The overall performance of the automatic typing is based on the quality of the reference dataset and the definition of coherent aerosol classes. For the reference dataset, well-characterized EARLINET data from Pappalardo et al. (2013), Papagiannopoulos et al. (2016), and Schwarz (2016) were used. The aerosol classification follows the procedure described in Section 2.1. The aerosol classes coincide with the typical values of the Section 2.2 (see Page 27 – Table 2).

The algorithm has been shown to be versatile and can be adjusted to the needs of each study. The reference dataset can be enlarged with well-characterized observations and increase the number of instances of under-represented aerosol classes. In addition, new aerosol classes can be added to describe other aerosol mixtures or the aerosol classes can be redefined. For example, an aerosol class Arabian dust can be inserted in the reference dataset in order to take into account of the different characteristics of the generating desert source.

The next paragraph is inserted in Section 2.2:

As an additional consideration, the defined aerosol types presented in Sect. 2.2 may not be representative of the entire aerosol load and, furthermore, apart from the dust mixtures they do not consider other aerosol mixtures. For example, this aspect can be observed in the definition of the volcanic category where the particles have different characteristics depending the transport pattern. The particles near the source have optical properties similar to desert dust whereas long-range transported volcanic plumes have the altered properties due to the sedimentation of the coarser particles. Therefore, it is important to further include a more exhaustive aerosol class analysis.

2. <u>Figure 2: Why do you show profiles below the full overlap of the lidar when you do not use them for your analysis? How trustworthy are the values in these height levels? What is meant by the statement 'the layers present the same behavior'? Looking at the profiles at different wavelengths I would suggest having</u>

different behaviors at different height levels, e.g. the wavelength dependence of the backscatter coefficient, the lidar ratio and the shape of the lidar ratio between 1.8 and 2 km is different to the height range between 2 and 3.6 km, above 3.6 km the Angstroem exponent of the extinction coefficient shows a different values than below.

The overlap height is at around 1,15 km a.s.l. for 1064 nm (Madonna et al., 2015). Values below the overlap region are not shown. We do not take into account values below the full overlap and the constructed database does not suffer from overlap issues whatsoever.

In the revised version of the paper we rephrased the sentence because it is actually misleading. The referee is right that three different layers are observed as reported in P5L19-20. The layer mean intensive parameters are given in Table A. The Ångström exponent for the 3 layers maintains a rather stable character with values around 0 suggesting large particles over Potenza. Regarding the lidar ratio, the values decrease with height and confirm the comment that we should not consider a single aerosol type throughout the range.

***Table B:** the mean intensive parameters for the 3 layers observed.*

Layer[km]	$\kappa_\beta(355,532)$	$\kappa_\beta(532,1064)$	$\kappa_\beta(355,1064)$	$\kappa_\alpha(355,532)$	S_{355} [sr]	S_{532} [sr]
1.6-2.0	0.45±0.03	0.32±0.03	0.37±0.02	-0.2±0.2	53±8	57±8
2.0-3.5	-0.02±0.12	0.42±0.06	0.26±0.04	-0.3±0.2	48±4	53±4
3.5-5.0	0.12±0.26	0.42±0.18	0.31±0.13	0.4±0.2	46±8	41±5

Besides, we investigated the backward trajectories for the layers of Table A using the HYSPLIT model (Stein et al., 2015). We initiated the model for a 7-day backward analysis and starting height levels the midpoints of the layers and the results are shown in Figure A. The layers 2.0-3.5 km and 3.5-5 km show a similar pattern with the air-masses flying over 2 km and originate from the Saharan desert. The layer 1.6-2.0 km follows a different pathway, the air-masses circulate over the Mediterranean Sea and Algeria and the uptake of marine particles is very likely. This information combined with the information of Table A suggests dust particles for the 2 elevated layers.

However, it is right to mention here that this case is reported for showing how the identification is done using intensive properties and backward trajectory analysis. All the intensive properties used for setting up the training and testing datasets are related to elevated layers separated from the planetary boundary layer and the local influence.

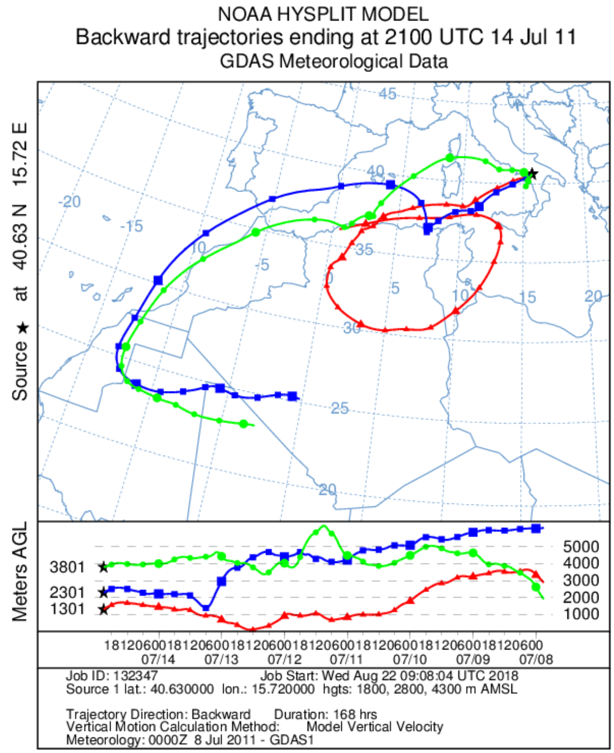

Figure J: *HYSPLIT backward trajectories for the aerosol layers observed over Potenza, on 14 July 2011, 19:20–22:10 UTC.*

In synthesis, the 3 layers with the synergistic use of intensive properties and backward trajectory analysis allows us to distinguish the different characteristics of the mixed dust layer in the range 1.6-2.0 km and the desert dust higher in the atmosphere (2.0-5.0 km). The intensive properties are measured with good level of uncertainty only in the layer 2.0-3.5 km. In order to avoid confusion, we will focus our comments in the range 2.0-3.5 km where the intensive property analysis is coherent and confirms the existence of a dust layer. The text implemented in P5L22-24 is given below:

The layer 2.0-3.5 km has a constant behavior with the range for the intensive optical profiles indicating the presence of the same type of particles. The mean values of all optical parameters in the range are calculated: lidar ratios of 48±4 sr at 355 nm and 53±4 sr at 532 nm and Ångström exponents (κ_β(355,1064), κ_β(532,1064), κ_β(355,532), κ_α(355,532)) of -0.3–0.4 were found.

3. Figure 3: Looking at the FLEXPART footprint, can you exclude a contribution from marine aerosols?

The manual typing as described in Section 2.1 is not a simple issue and, of course, leaves room to questionable type assignment. The backward trajectory analysis is used synergistically with the lidar optical properties and the model outputs suffer from the high error on the path (increasing with the path itself) and the source term assignment. Therefore, the model simulations have to be used together with the observed lidar optical properties for providing a reasonable aerosol typing.

The FLEXPART seems to indicate the possibility of marine influence in the identified layer, however the lidar ratio values being over 50 sr indicates low or no presence of maritime aerosol particles. Furthermore, for this specific case, particle linear depolarization ratio measurements are available, however these measurements are not *included in the database and therefore not reported in the submitted manuscript. Figure B shows the particle linear depolarization ratio where the values are over 0.2 and, thus, confirming our hypothesis of aspherical particles.*

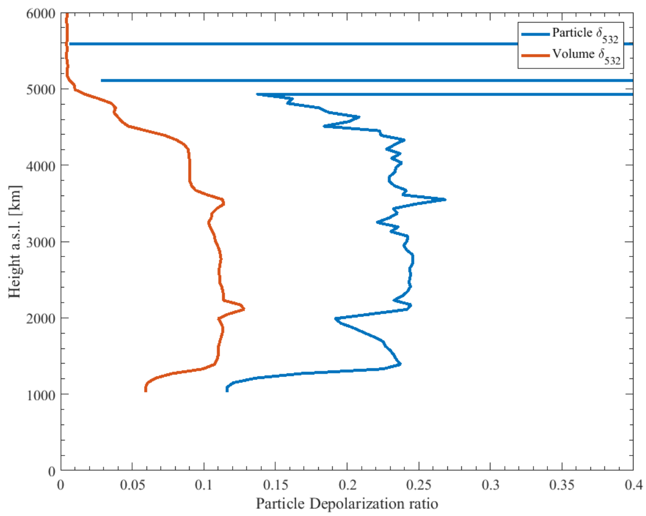

Figure K: *Particle (blue line) and Volume (red line) linear depolarization ratio at 532 nm for the MUSA system in Potenza on 14/07/2011, 19:20–22:10 UTC.*

In order to avoid confusion, we rephrased the sentence following the referee's comment in P5L30-31 and moved P5L24-25 to the end of it:

The dust-prone area of northern Africa (Morocco and northern Algeria) along with the Mediterranean Sea are most likely the sources of the observed layer and suggest a mixture of dust and marine particles. The combined information of the backward trajectory analysis and the intensive properties values indicate the presence of dust particles and they are in accordance with the typical dust values observed over Potenza (Mona et al., 2014).

4. <u>Figure 5: An additional Figure also including the information of the depolarization ratio for the different classes would be valuable.</u>

Figure C shows the characteristics of the reference dataset in terms of the depolarization ratio when plotted against the lidar ratio at 532 nm and the backscatter related Ångström exponent (355,1064) for 8 aerosol classes. The particle depolarization ratio values were assigned to the aerosol classes assuming that the intensive properties are normally distributed and using literature values as reported in P29-Table 4. Therefore, the values have no standard deviation. The figures highlight the discriminatory power of the classifying parameter among the dust-like aerosol classes, however, the particle depolarization ratio seems to have no power to separate the non-dust classes.

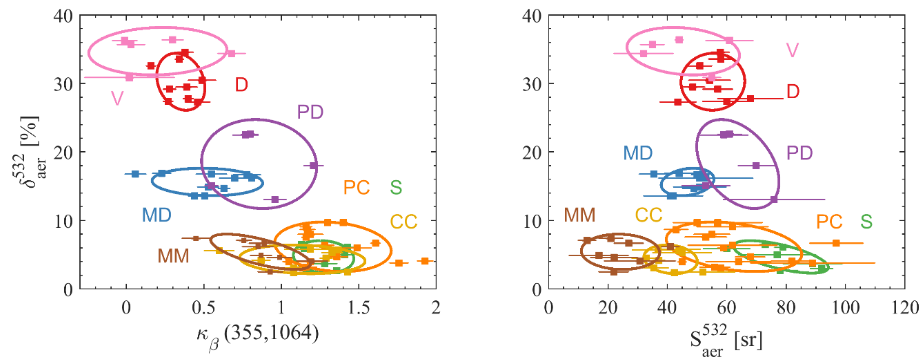

***Figure L**: Colored pre-specified classes and 90 % confidence ellipses for 8 aerosol classes using the classifying parameters: δ_{532}, S_{532}, κ_β(355,1064). The error bars correspond to the standard deviation of the selected mean intensive properties. CC stands for Clean Continental, D stands for dust, MD stands for mixed dust, MM stands for mixed marine, PD stands for polluted dust, PC stands for polluted continental, S stands for smoke, and V stands for volcanic particles.*

In conjunction with the specific comment #9 of RC1, we inserted in the revised manuscript Figure D and the text below in Section 3.2.4:

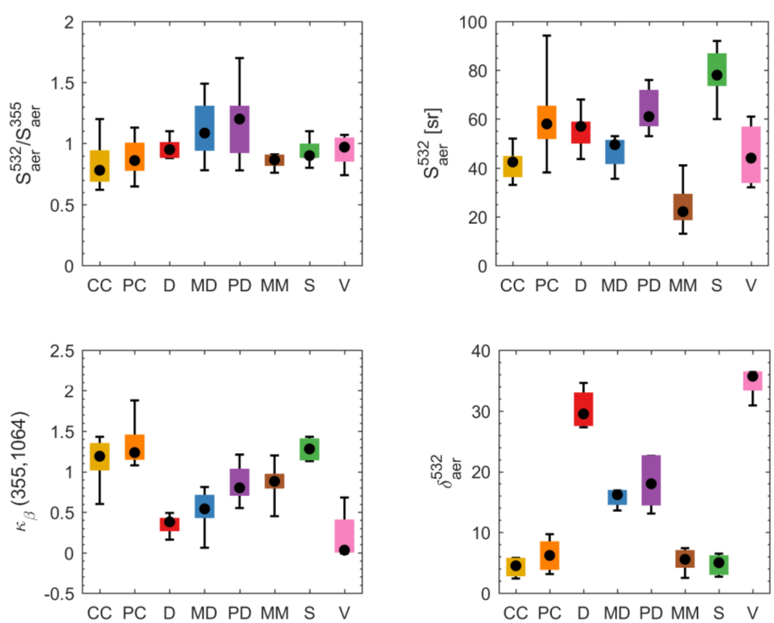

***Figure M:** Bar plots show the median (horizontal line), 25-75 percentile (box) and 5-95 percentile (whisker) of the four classifying parameters: δ_{532}, κ_β(355;1064), S532, and S_{532}/S_{355}. CC stands for Clean Continental, D stands for dust, MD stands for mixed dust, MM stands for mixed marine, PD stands for polluted dust, PC stands for polluted continental, S stands for smoke, and V stands for volcanic particles.*

Figure 7 presents cumulative <u>bar plots</u> with the median (black dots), the 25-75 percentile (box), the 5-95 percentile (whiskers) for all 4 classifying parameters. The figure highlights the discriminatory power of δ_{532}, $\kappa_\beta(355,1064)$, and S_{532}, whereas the S_{532}/S_{355} performs the worst. Furthermore, the figure depicts the discriminatory power of the classifying parameter among the dust-like aerosol classes, however, the particle depolarization ratio seems to have no power to separate the non-dust classes, as discussed above.

5. <u>Figure 7: The shape of the backscatter coefficient and the extinction coefficient at 355 and at 532 nm show different shapes, but the derived profile of the lidar ratio for both wavelengths shows the same shape. What is the vertical resolution of the different profiles? Did the extinction and backscatter coefficient have the same resolution for deriving the lidar ratio?</u>

P34-Figure 2 and P39-Figure 7 in the submitted paper are reporting the profiles at their highest resolutions, i.e. the particle backscatter coefficient profiles have a higher resolution when compared to the particle extinction coefficient ones and typically ultraviolet profiles have a better resolution than visible profiles. The lidar ratio profiles have the same resolution of the particle extinction coefficient profiles. Prior to the calculation of the lidar ratio, the particle backscatter coefficient profiles are smoothed using a 2nd order Savitzky-Golay filter at an effective vertical resolution that varies with height, for more details see Iarlori et al. (2015).

In particular, for P34-Figure 7, the vertical raw resolution of the particle backscatter coefficient profiles is 7.5 m while for the extinction coefficient the resolution varies with height. The effective resolution for 355 nm (blue line) and 532 nm (red line) is given in Figure E and follows the procedure described in Pappalardo et al. (2004). For the height range 1.2-1.9 km, the highest resolution is 240 m for 355 nm while for 532 nm varies from 240 m to 480 m. Higher in the atmosphere, the resolution degrades with height (faster for 532 nm) and becomes constant at 4.7 km for 355 nm (3.8 km for 532 nm).

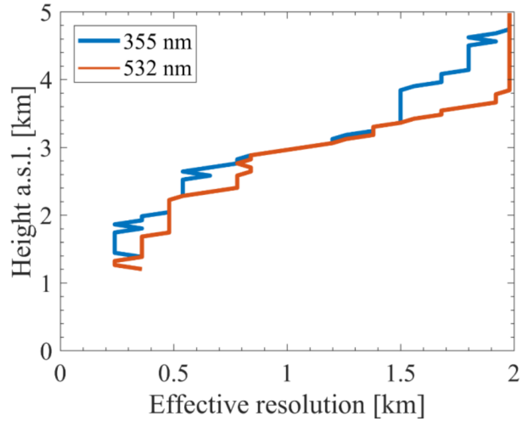

Figure N: *The effective resolution of the extinction coefficient at 355 nm (blue line) and 532 nm (red line) for the Athens lidar on 22/05/2014, 20:28–21:28 UTC.*

Figure F shows the same panels as for P34-Figure 7 along with the smoothed backscatter profiles (panel b) that were used for the calculation of the lidar ratio. The next sentence is implemented in the revised manuscript:

The effective resolution of the extinction coefficient profiles varied from 240 m to 780 m using the method described in Pappalardo et al. (2004).

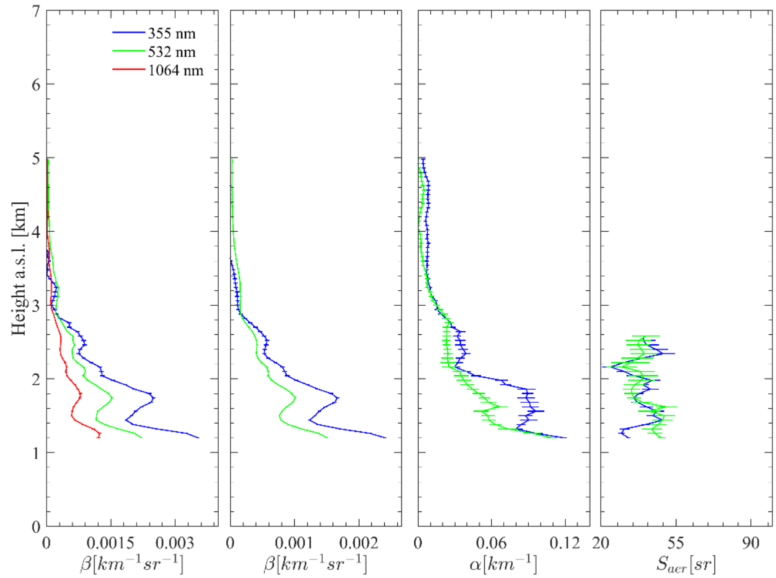

Figure O: *Optical profiles measured in Athens, on 22 May 2014, 20:28–21:28 UTC with a multiwavelength Raman lidar. From left to right, (a) the backscatter coefficient with the full resolution, (b) the smoothed backscatter coefficient, (c) the extinction coefficient, and (d) the lidar ratio. The error bars correspond to the standard deviation.*

References

Alados-Arboledas, L., Müller, D., Guerrero-Rascado, J. L., Navas-Guzmán, F., Pérez-Ramírez, D., and Olmo, F. J.: Optical and microphysical properties of fresh biomass burning aerosol retrieved by Raman lidar, and star- and sun-photometry, Geophys. Res. Lett., 38, doi:10.1029/2010GL045999, 2011.

Amiridis, A., Balis, D. S., Giannakaki, E., Stohl, A., Kazadzis, S., Koukouli, M. E., and Zanis, P.: Optical characteristics of biomass burning aerosols over Southeastern Europe determined from UV-Raman lidar measurements, Atm. Chem. and Phys., 9, 2431–2440, doi:10.5194/acp-9-2431-2009, 2009.

Baars, H., Ansmann, A., Althausen, D., Engelmann, R., Heese, B., Müller, D., Artaxo, P., Paixao, M., Pauliquevis, M., and Souza, R.: Aerosol profiling with lidar in the Amazon Basin during the wet and dry season, J. Geophys. Res, 117, doi:10.1029/2012JD018338, 2012.

Baars, H., Seifert, P., Engelmann, R., and Wandinger, U.: Target categorization of aerosol and clouds by continuous multiwavelength-polarization lidar measurements, Atmos. Meas. Tech., 10, 3175-3201, https://doi.org/10.5194/amt-10-3175-2017, 2017.

Burton, S. P., Ferrare, R. A., Hostetler, C. A., Hair, J. W., Rogers, R. R., Obland, M. D., Butler, C. F., Cook, A. L., Harper, D. B., and Froyd, K. D.: Aerosol classification using airborne High Spectral Resolution Lidar measurements–methodology and examples, Atmos. Meas. Tech., 5, 73–98, doi:10.5194/amt-5-73-2012, 2012.

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.

Burton, S. P., Vaughan, M. A., Ferrare, R. A., and Hostetler, C. A.: Separating mixtures of aerosol types in airborne High Spectral Resolution Lidar data, Atmos. Meas. Tech., 7, 419–436, doi:10.5194/amt-7-419-2014, 2014.

Giannakaki, E., van Zyl, P. G., Müller, D., Balis, D., and Komppula, M.: Optical and microphysical characterization of aerosol layers over South Africa by means of multi-wavelength depolarization and Raman lidar measurements, Atmos. Chem. Phys., 16, 8109–8123, doi:10.5194/acp-16-8109-2016, 2016.

Groß, S., Esselborn, M., Weinzierl, B., Wirth, M., Fix, A., and Petzold, A.: Aerosol classification by airborne high spectral resolution lidar observations, Atmospheric Chemistry and Physics, 13, 2487–2505, doi:10.5194/acp-13-2487-2013, 2013.

Iarlori, M., Madonna, F., Rizi, V., Trickl, T., and Amodeo, A.: Effective resolution concepts for lidar observations, Atmos. Meas. Tech., 8, 5157-5176, https://doi.org/10.5194/amt-8-5157-2015, 2015.

Kahn, R. A., and Gaitley, B. J.: An analysis of global aerosol type as retrieved by MISR, J. Geophys. Res. Atmos., 120, 4248–4281. doi:10.1002/2015JD023322, 2015.

Kim, M.-H., Kim, S.-W., Yoon, S.-C., and Omar, A. H.: Comparison of aerosol optical depth between CALIOP and MODIS-Aqua for CALIOP aerosol subtypes over the ocean, J. Geophys. Res.-Atmos., 118, 13241–13252, doi:10.1002/2013JD019527, 2013.

Madonna, F., Amato, F., Vande Hey, J., and Pappalardo, G.: Ceilometer aerosol profiling versus Raman lidar in the frame of the INTERACT campaign of ACTRIS, Atmos. Meas. Tech., 8, 2207-2223, https://doi.org/10.5194/amt-8-2207-2015, 2015.

Mona, L., Papagiannopoulos, N., Basart, S., Baldasano, J., Binietoglou, I., Cornacchia, C., and Pappalardo, G.: EARLINET dust observations vs. BSC-DREAM8b modeled profiles: 12-year-long systematic comparison at Potenza, Italy, Atmospheric Chemistry and Physics, 14, 8781–8793, doi:10.5194/acp-14-8781-2014, 2014.

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

Nepomuceno Pereira, S., Preißler, J., Guerrero-Rascado, J. L., Silva, A. M., and Wagner, F.: Forest Fire Smoke Layers Observed in the Free Troposphere over Portugal with a Multiwavelength Raman Lidar: Optical and Microphysical Properties, Scientific World Journal, 2014, doi:10.1155/2014/42183, 2014.

Nicolae, D., Nemuc, A., Müller, D., Talianu, C., Vasilescu, J., Belegante, L., and Kolgotin, A.: Characterization of fresh and aged biomass burning events using multiwavelength Raman lidar and mass spectrometry, J. Geophys. Res. Atmos., 118, 2956–2965, doi:10.1002/jgrd.50324, 2013.

Ortiz-Amezcua, P., Guerrero-Rascado, J. L., Granados-Muñoz, M. J., Benavent-Oltra, J. A., Böckmann, C., Samaras, S., Stachlewska, I. S., Janicka, Ł., Baars, H., Bohlmann, S., and Alados-Arboledas, L.: Microphysical characterization of long-range transported biomass burning particles from North America at three EARLINET stations, Atmospheric Chemistry and Physics, 17, 5931–5946, doi:10.5194/acp-17-5931-2017, 2017.

Papagiannopoulos, N., Mona, L., Alados-Arboledas, L., Amiridis, V., Baars, H., Binietoglou, I., Bortoli, D., D'Amico, G., Giunta, A., Guerrero-Rascado, J. L., Schwarz, A., Perreira, S., Spinelli, N., Wandinger, U., Wang, X., and Pappalardo, G.: CALIPSO climatological products: evaluation and suggestions from EARLINET, Atmos. Chem. Phys., 16, 2341–2357, doi:10.5194/acp-16-2341-2016, 2016a.

Papayannis, A., Argyrouli, A., Müller, D., Tsaknakis, G., Kokkalis, P., Binietoglou, I., Kazadzis, S., Solomos, S., and Amiridis, V.: Tropospheric Vertical Profiles of Aerosol Optical, Microphysical and Concentration Properties in the Frame of the Hygra-CD Campaign (Athens, Greece 2014): A Case Study of Long-Range Transport of Mixed Aerosols, EPJ Web of Conferences, 119, 23 016, doi:10.1051/epjconf/201611923016, 2016.

Pappalardo, G., Amodeo, A., Pandolfi, M., Wandinger, U., Ansmann, A., Bösenberg, J., Matthias, V., Amiridis, V., De Tomasi, F., Frioud, M., Iarlori, M., Komguem, L., Papayannis, A., Rocadenbosch, F., and Wang, X.: Aerosol lidar intercomparison in the framework of the EARLINET project. 3. Raman lidar algorithm for aerosol extinction, backscatter, and lidar ratio, Appl. Opt., 43, 5370-5385. 2004.

Pappalardo, G., Mona, L., D'Amico, G., Wandinger, U., Adam, M., Amodeo, A., Ansmann, A., Apituley, A., Alados Arboledas, L., Balis, D., Boselli, A., Bravo-Aranda, J. A., Chaikovsky, A., Comeron, A., Cuesta, J., De Tomasi, F., Freudenthaler, V., Gausa, M., Giannakaki, E., Giehl, H., Giunta, A., Grigorov, I., Groß, S., Haeffelin, M., Hiebsch, A., Iarlori, M., Lange, D., Linné, H., Madonna, F., Mattis, I., Mamouri, R.-E., McAuliffe, M. A. P., Mitev, V., Molero, F., Navas-Guzman, F., Nicolae, D., Papayannis, A., Perrone, M. R., Pietras, C., Pietruczuk, A., Pisani, G., Preißler, J., Pujadas, M., Rizi, V., Ruth, A. A., Schmidt, J., Schnell, F., Seifert, P., Serikov, I., Sicard, M., Simeonov, V., Spinelli, N., Stebel, K., Tesche, M., Trickl, T., Wang, X., Wagner, F., Wiegner, M., and Wilson, K. M.: Four-dimensional distribution of the 2010 Eyjafjallajökull volcanic cloud over Europe observed by EARLINET, Atmos. Chem. Phys., 13, 4429-4450, https://doi.org/10.5194/acp-13-4429-2013, 2013.

Pappalardo, G., Amodeo, A., Apituley, A., Comerón, A., Freudenthaler, V., Linné, H., Ansmann, A., Bösenberg, J., D'Amico, G., Mattis, I., Mona, L., Wandinger, U., Amiridis, V., Alados-Arboledas, L., Nicolae, D., and Wiegner, M.: EARLINET: towards an advanced sustainable European aerosol lidar network, Atmos. Meas. Tech., 7, 2389–2409, doi:10.5194/amt-7-2389-2014, 2014.

Rogers, R. R., Vaughan, M. A., Hostetler, C. A., Burton, S. P., Ferrare, R. A., Young, S. A., Hair, J. W., Obland, M. D., Harper, D. B., Cook, A. L., and Winker, D. M.: Looking through the haze: evaluating the CALIPSO level 2 aerosol optical depth using airborne high spectral resolution lidar data, Atmos. Meas. Tech., 7, 4317-4340, https://doi.org/10.5194/amt-7-4317-2014, 2014.

Schwarz, A.: Aerosol typing over Europe and its benefits for the CALIPSO and EarthCARE missions - statistical analysis based on multiwavelength aerosol lidar measurements from ground-based EARLINET stations and comparison to spaceborne CALIPSO data, Ph.D. thesis, University of Leipzig, 2016.

Stein, A.F., Draxler, R.R, Rolph, G.D., Stunder, B.J.B., Cohen, M.D., and Ngan, F.: NOAA's HYSPLIT atmospheric transport and dispersion modeling system, Bull. Amer. Meteor. Soc., 96, 2059-2077, doi:10.1175/BAMS-D-14-00110.1, 2015.

Wang, X., Boselli, A., D'Avino, L., Pisani, G., Spinelli, N., Amodeo, A., Chaikovsky, A., Wiegner, M., Nickovic, S., Papayannis, A., Perrone, M., Rizi, V., Sauvage, L., and Stohl, A.: Volcanic dust characterization by EARLINET during Etna's eruptions in 2001-2002, Atmospheric Environment, 42, 893–905, doi:10.1016/j.atmosenv.2007.10.020, 2008.

Wilks, D. S.: Statistical methods in the atmospheric sciences, San Diego, Academic Press, 2006.

An automatic observation-based aerosol typing method for EARLINET

Nikolaos Papagiannopoulos[1,2], Lucia Mona[1], Aldo Amodeo[1], Giuseppe D'Amico[1], Pilar Gumà Claramunt[1], Gelsomina Pappalardo[1], Lucas Alados-Arboledas[3,4], Juan Luís Guerrero-Rascado[3,4], Vassilis Amiridis[5], Arnoud Apituley[6], Holger Baars[7], Anja Schwarz[7], Ulla Wandinger[7], Ioannis Binietoglou[8], Doina Nicolae[8], Daniele Bortoli[9], Adolfo Comerón[2], Alejandro Rodríguez-Gómez[2], Michaël Sicard[2,10], Panagiotis Kokkalis[11,12], Alex Papayannis[12], and Matthias Wiegner[13]

[1]Consiglio Nazionale delle Ricerche, Istituto di Metodologie per l'Analisi Ambientale (CNR-IMAA), C.da S. Loja, Tito Scalo (PZ), 85050, Italy
[2]CommSensLab, Dept. of Signal Theory and Communications, Universitat Politècnica de Catalunya, Barcelona, Spain
[3]Andalusian Institute for Earth System Research (IISTA-CEAMA), 18006, Granada, Spain
[4]Department of Applied Physics, University of Granada, 18071, Granada, Spain
[5]IAASARS, National Observatory of Athens, Athens, Greece
[6]Royal Netherlands Meteorological Institute KNMI, De Bilt, the Netherlands
[7]Leibniz Institute for Tropospheric Research (TROPOS), Leipzig, Germany
[8]National Institute of R&D for Optoelectronics (INOE), Magurele, Romania
[9]Earth Science Institute-(ICT), Évora, Portugal
[10]Ciències i Tecnologies de l'Espai - Centre de Recerca de l'Aeronàutica i de l'Espai/Institut d'Estudis Espacials de Catalunya (CTE-CRAE/IEEC), Universitat Politècnica de Catalunya, Barcelona, Spain
[11]Laser Remote Sensing Unit, Physics Dept., National Technical University of Athens, Athens, Greece
[12]Physics Department, Faculty of Science, Kuwait University, Kuwait
[13]Ludwig-Maximilians-Universität (LMU), Meteorologisches Institut, Theresienstraße 37, 80333 Munich, Germany

Correspondence to: Nikos Papagiannopoulos (nikolaos.papagiannopoulos@imaa.cnr.it)

Abstract. We present an automatic aerosol classification method based solely on European Aerosol Research Lidar Network (EARLINET) intensive optical parameters with the aim of building a network-wide classification tool that could provide near-real-time aerosol typing information. The presented method depends on a supervised learning technique and makes use of the Mahalanobis distance function that relates each un-classified measurement to a pre-defined aerosol type. As a first step (training phase), a reference dataset is set up consisting of already classified EARLINET data. Using this dataset, we defined eight aerosol classes: clean continental, polluted continental, dust, mixed dust, polluted dust, mixed marine, smoke, and volcanic ash. The effect of the number of aerosol classes has been explored, as well as the optimal set of intensive parameters to separate different aerosol types. Furthermore, the algorithm is trained with literature particle linear depolarization ratio values. As a second step (testing phase), we apply the method to an already classified EARLINET dataset and analyse the results of the comparison to this classified dataset. The predictive accuracy of the automatic classification varies between 59 % (minimum) and 90 % (maximum) from 8 to 4 aerosol classes, respectively, when evaluated against pre-classified EARLINET lidar. This indicates the potential use of the automatic classification to all network lidar data. Furthermore, the training of the algorithm with particle linear depolarization values found in literature further improves the accuracy: the accuracy range is 69–93 % from

8 (minimum) to 4 (maximum) aerosol classes, respectively. Additionally, the algorithm has proven to be highly versatile as it adapts to changes in the size of the training dataset and the number of aerosol classes and classifying parameters. Finally, the low computational time and demand for resources make the algorithm extremely suitable for the implementation within the Single Calculus Chain (SCC), the EARLINET centralised processing suite.

1 Introduction

The European Aerosol Research Lidar Network (EARLINET; Pappalardo et al., 2014) operates Raman lidars at a continental scale. Since the beginning, the network aimed towards a sustainable observing system that has been achieved by developing a quality assurance strategy, and optimizing instruments and data. To this direction and towards future advancement, the network plans continuous measurements and near-real-time data delivery. With this in mind, the Single Calculus Chain (SCC; D'Amico et al., 2015) for the automatic lidar analysis has been developed and currently delivers profiles of optical aerosol properties. The EARLINET SCC explores the implementation of new features like profiles of intensive optical properties and determination of aerosol layer geometrical properties. The intensive optical properties are type-dependent and can be used to classify the observed layers into aerosol types. The categorization into different types provides significant help to understand aerosol sources, their effects, and feedback mechanisms to improve the accuracy of satellite retrievals and to quantify assessments of aerosol radiative impacts on climate (Russell et al., 2014) by intercomparing numerical models such as NWP (Numerical Weather Prediction) and CTM (Chemical Transport Model) (Baklanov et al., 2014). Thus, EARLINET by providing multi-wavelength range resolved aerosol properties, has an added value for aerosol typing. In this study we present a flexible automatic method to classify EARLINET data.

Lidar systems are capable of identifying multiple layers in the atmosphere owing to their high vertical resolution (on the orders of tens of meters). Thus, lidar-based retrievals can provide a separate classification for each layer and are not confined to columnar classifications as in the case of sunphotometers. The lidar technique has proven to be a robust tool to classify aerosols with its capability of polarization-sensitive and multi-wavelength measurements (Liu et al., 2008). Sophisticated lidars, such as the High Spectral Resolution Lidars (HSRL) and the multi-wavelength Raman lidars, offer a multitude of intensive parameters that characterize different aerosol types (e.g., Müller et al., 2007a; Burton et al., 2012; Groß et al., 2013). Typically, the particle extinction-to-backscatter ratio (i.e., particle lidar ratio), the particle linear depolarization ratio at one or more wavelengths and the wavelength dependence of extinction and/or backscatter coefficients (i.e., extinction- or backscatter-related Ångström exponents) are considered.

The increasing amount of available information and particularly the plethora of lidar intensive parameters, can offer a more accurate aerosol classification as well as insight into the various aerosol types (Burton et al., 2013). Consequently, objective, multivariate analysis is needed to take advantage of this information. Automatic algorithms are, therefore, employed to classify aerosol into respective types. These procedures make use of various classifiers that are able to quantify the differences between the aerosol classes. In classification analysis, the observations are allocated to a known number of groups, that is, a supervised learning technique. Whereas in cluster analysis, ~~neither the number of groups nor the groups are~~ the groups are not known beforehand and the classifier is tasked with it.

The measured values are evaluated by the classification function to find the group to which the individual most likely belongs. Specifically, distance-based classification techniques (e.g., k- nearest neighbour, support vector machine algorithms) are straightforward, that is, the classification depends on the distance from the target instance to the training instance. The Mahalanobis distance classifier (Mahalanobis, 1936) has a wide range of applications and can be used to categorize data points, each representing an observation, into classes that have predefined characteristics. The distances between the observation and the different classes are calculated, and then the observation is attributed to the class for which the distance is the minimum.

The Mahalanobis-distance-based classification found great applicability in aerosol studies. For instance, the algorithm developed by Burton et al. (2012) makes use of four lidar intensive properties, namely the particle linear depolarization ratio at $532\,\mathrm{nm}$, the particle lidar ratio at $532\,\mathrm{nm}$, the backscatter-related 532-to-1064- nm color ratio, and the ratio of particle linear depolarization ratios at 1064 and $532\,\mathrm{nm}$ in order to classify aerosols into eight types. A slightly different algorithm including also the uncertainties on the input properties was introduced by Russell et al. (2014). Their algorithm was applied to satellite derived optical and physical data. The reference dataset was obtained from AERONET (Aerosol Robotic Network; Holben et al., 1998) stations where a single aerosol type tends to dominate (e.g., Cattrall et al., 2005). The pre-specified classes were then applied to a five-year record of retrievals from the spaceborne POLDER 3 (Polarization and Directionality of the Earth's Reflectances 3; Tanré et al., 2011) polarimeter on PARASOL (Polarization and Anisotropy of Reflectances for Atmospheric Sciences coupled with Observations from a Lidar; Tanré et al., 2011) spacecraft. Recently, Hamill et al. (2016) used the same classifier to produce an aerosol classification scheme based on long term AERONET data.

In this work, we present a method analogous to the one proposed by Burton et al. (2012), modified to fit EARLINET's needs and capabilities. The aerosol typing exclusively makes use of EARLINET lidar-derived intensive property data. We use the Mahalanobis distance as a classifier to assign any given multi-dimensional observation to the pre-specified aerosol class to which it is most similar. These classes are defined using an EARLINET-based classification scheme. The EARLINET classification scheme is presented in Sect. 2 where we also describe the parameters readily delivered by the network that can be used to classify aerosols. Furthermore, the major aerosol types that comprise the aerosol classes onto which the aerosol classification is based are presented. In Sect. 3 the method that we apply to EARLINET data is explained, and, we present the training phase. We set up a scheme for investigating the number of aerosol classes and we perform an analysis to identify the intensive parameters that contribute the most to the classification as well. Sect. 4 describes the testing phase and provides a discussion of the results of the classification. The paper closes with conclusions of our study and suggestions for further applications and improvements.

2 Operational network–EARLINET

EARLINET (www.earlinet.org) was established in 2000, providing aerosol profiling data on a continental scale, and now is part of the Aerosols, Clouds, and Trace gases Research InfraStructure (ACTRIS; www.actris.eu/). In these 18 years of continuous existence, EARLINET has evolved both in the number of contributing stations, as well as in its observing capacity (Pappalardo et al., 2014). Currently, 30 stations are submitting aerosol extinction and/or backscatter coefficient profiles to the EARLINET database, according to EARLINET's measurement schedule (one daytime and two nighttime measurements per week). Therefore these systematic observations consolidate a 4D European quantitative and statistically significant aerosol survey. Further measurements are devoted to special events, such as volcanic eruptions, forest fires, and desert dust outbreaks. Moreover, EARLINET provides correlative measurements during CALIPSO (Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations) overpasses on each EARLINET station in order to validate satellite products (e.g., Mamouri et al., 2009; Mona et al., 2009). Throughout the paper, we refer to measurements as a set of aerosol optical profiles reported in the EARLINET database that correspond to the same temporal window, which typically extends for about 1 hour.

The majority of the EARLINET stations (67 % of the stations; Pappalardo et al., 2014) ~~operates~~ operate multi-wavelength Raman lidars that combine a set of elastic and nitrogen inelastic channels, typically consisting of three elastic and two inelastic Raman channels (the so-called $3\beta + 2\alpha$ configuration). In particular, they provide the aerosol extinction (at 355 nm and 532 nm), and backscatter coefficients (at 355 nm, 532 nm, and 1064 nm). This configuration allows the retrieval of the range-resolved particle lidar ratio at 355 nm and 532 nm – S^{λ}_{aer}. This intensive parameter depends on the shape, size, and chemical composition of the aerosol (Müller et al., 2007a). When lidar ratio is available for more than one wavelength, the corresponding color ratio can be also retrieved – $S^{\lambda_1}_{aer}/S^{\lambda_2}_{aer}$. This quantity ~~is a robust means~~ has shown the ability to characterize the ageing status of smoke particles as well as the spectral dependence of aerosol ~~(Müller et al., 2007a; Alados-Arboledas et al., 2011; Nicolae et al., 2013; Nepomuceno Pereira et al., 2014; Samaras et al., 2015)~~ (Müller et al., 2007a; Alados-Arboledas et al., 2011; Nicolae et al., 2013; Nepomuceno Pereira et al., 2014) The combination of the optical data allows the retrieval of the size sensitive backscatter and/or extinction related Ångström exponent and can be calculated as

$$\kappa_X(\lambda_1, \lambda_2) = \frac{ln[X(\lambda_1)/X(\lambda_2)]}{ln(\lambda_2/\lambda_1)} \tag{1}$$

with X denoting the backscatter (β) or extinction coefficient (α) for a set of wavelengths, λ_1 and λ_2. Moreover, 52 % of EARLINET stations (Pappalardo et al., 2014) are equipped with depolarization channels, thus providing profiles of the particle linear depolarization ratio. It can be calculated according to Biele et al. (2000); Freudenthaler et al. (2009)

$$\delta^{\lambda}_{aer} = \frac{(1+\delta_m)\delta_v R - (1+\delta_v)\delta_m}{(1+\delta_m)R - (1+\delta_v)} \tag{2}$$

with R the backscatter ratio, δ_m the molecular depolarization, and δ_v the volume depolarization ratio. This parameter provides information on the particle shape, thus enhancing the aerosol typing strength of the network. Under favourable conditions, the aerosol microphysical properties, such as the effective radius, the volume concentration and the refractive index can also be retrieved through complex numerical algorithms (e.g., Müller et al., 2004; Veselovskii et al., 2010; Bovchaliuk et al., 2016; Chaikovsky et al., 2016).

The data products described above make the EARLINET data an excellent basis to perform aerosol typing at continental scale. Examples of methodologies to classify aerosol datasets can be found in e.g., Müller et al. (2007a, b); Groß et al. (2011);

Mona et al. (2012a); Navas-Guzmán et al. (2013); Baars et al. (2016). At the time being there are different algorithms under development which combine measurements and aerosol models (Nicolae et al., 2016; Wandinger et al., 2016). Nevertheless, automated observation-based algorithms working at network level for the identification of layers, their boundaries, and the corresponding aerosol typing are not yet available. The SCC tool for automatic processing of EARLINET lidar signals is, currently, providing primarily profiles of particle extinction and backscatter coefficients and volume and particle depolarization ratio. The SCC aims at incorporating modules for layer identification, intensive properties retrieval, and aerosol typing. Therefore, this paper could provide a starting point for a harmonized EARLINET classification tool that could also be used by other lidar networks, like the ones involved in GALION (GAW Aerosol Lidar Observation Network), the GAW (Global Aerosol Watch) initiative for the aerosol lidar observation on a global scale, and within aerosol lidar studies in general.

2.1 EARLINET manual aerosol classification

The typical procedure for aerosol categorization adopted within the EARLINET community consists of three main steps:

1. layer identification and cloud screening,
2. identification of the geometrical properties (boundaries, center of mass) of the aerosol layer, and
3. the aerosol layer typing by means of investigation of intensive optical properties (Ångström exponents, lidar ratios, and particle linear depolarization ratios), model outputs (backward trajectory analyses), and ancillary instruments data if available (e.g., satellite or sunphotometer data).

In what follows, an example of an aerosol type assignment using EARLINET data is presented. Figure 1 shows the temporal evolution of the range corrected signal at 1064 nm from a measurement made in Potenza, Italy, on 14 July 2011, 19:20–22:10 UTC with the reference lidar system MUSA (Multiwavelength system for aerosol) of CNR-IMAA (Consiglio Nazionale delle Ricerche - Istituto di Metodologie per l'Analisi Ambientale). High values show a stratified aerosol layer from the ground up to 5 km, whereas low values indicate aerosol free regions. The lowest altitude range presents the overlap between the laser beam and the receiver field of view and, therefore, it is the blind range of the lidar. MUSA has a full overlap at around 1.15 km a.s.l. for 1064 nm (Madonna et al., 2015). The optical thicker layer lies below 2 km, with a distinct layer atop extending up to 3.5 km and, finally, an optically thinner region from 3.5 km to 5 km. The retrieved profiles for the same temporal window of particle backscatter and extinction coefficient, lidar ratios, and Ångström exponents are shown in Fig. 2. The ~~aforementioned layers present the same behavior as seen in the~~ particle extinction and backscatter coefficient are given with their full resolution. To calculate the lidar ratio, the backscatter coefficient was smoothed in the same effective vertical resolution using a Savitzky-Golay second order filter (Iarlori et al., 2015) and only the useful range of signals was kept; the effective resolution of the resulting profiles varied from 120 m to 480 m using the method described in Pappalardo et al. (2004b). The layer 2.0-3.5 km has a constant behavior with the range for the intensive optical profiles ~~and, thus, correspond to the~~ indicating the presence of the same type of particles. The mean values of all optical parameters in the range ~~1.6-5~~ are calculated: lidar ratios of ~~53 ± 7~~48 ± 4 sr at 355 nm and ~~55 ± 8~~53 ± 4 sr at 532 nm and Ångström exponents ~~of~~

~~0–0~~(i.e., $\kappa_\beta(355,1064)$, $\kappa_\beta(532,1064)$, $\kappa_\beta(355,532)$, and $\kappa_\alpha(355,532)$) of -0.3–0.4 were found. ~~These values indicate the presence of coarse particles and they are in accordance with the typical dust values observed over Potenza (Mona et al., 2014) .~~

For the classification of aerosols with respect to their source regions and age, auxiliary information like results of transport and dispersion models or satellite data are used. For the observed aerosol layer, the Lagrangian dispersion model FLEXPART (FLEXible PARTicle dispersion model; Stohl et al., 2005) model was used for a 5-day backward simulation. Figure 3 shows the so-called footprint that indicates the areas of the air parcels travelling below 2 km before reaching the study area. The model output is given in terms of the decimal logarithm of the integrated residence time in seconds in a grid box. The most probable aerosol source region and the aerosol type were assigned accordingly. The dust-prone area of northern Africa (Morocco and northern Algeria) ~~is~~ along with the Mediterranean Sea are most likely the ~~source of the particles and in conjunction with the lidar-derived information the inferred type is pure dust~~ sources of the observed layer and suggest a mixture of dust and marine particles. The combined information of the backward trajectory analysis and the intensive properties values indicate the presence of dust particles and they are in accordance with the typical dust values observed over Potenza (Mona et al., 2014).

In the following, the characteristics of the major aerosol types are presented. These aerosol types are used for the automatic classification and correspond to aerosol layers typically encountered over Europe.

2.2 Aerosol Types

One of the defining characteristics of the aerosol properties is the source; aerosols found in the atmosphere can be, for example, mineral particles from arid areas of the Earth or organic carbon emitted during biomass burning events. Due to the multiple influence of the aerosol origin on the properties, aerosol sources can be used to classify them into different categories. In this section, we provide an overview of the main aerosol types observed over the EARLINET stations followed by the corresponding optical properties. This section also aims to provide the important information of the aerosol types that the automatic classification is based upon. The considered aerosol types almost coincide with the ones used in the CALIPSO classification scheme (Omar et al., 2009) which provides already a satisfactory description of the atmospheric aerosol content. Moreover, adopting similar classification schemes, the direct comparison of the proposed typing against the CALIPSO product is possible.

2.2.1 Continental

Man-made activities dictate the aerosol pattern within the atmospheric boundary layer, and affect the observations in the lower troposphere in Europe. Anthropogenic particles show a strong wavelength dependence of their optical properties, i.e., high Ångström exponent values. Moreover, they are typically small and do not significantly depolarize the backscattered light ($\delta_{aer}^{532} = 0.04 \pm 0.04$; Heese et al., 2016), and due to the high carbon content, these particles reveal high lidar ratios (Giannakaki et al., 2010). We refer to this particle type, herein, as polluted continental.

Typically, the clean continental aerosol over Europe is a mixture of anthropogenic pollution with particles from natural sources. The clean continental type shows low depolarizing ability with values lower than 0.07 (Omar et al., 2009), low lidar ratio values, i.e., 20–40 sr and relatively high Ångström exponents, i.e., 1.0–2.5 (Ansmann et al., 2001; Giannakaki et al., 2010).

The clean continental, therefore, differentiates from the polluted continental type due to ~~the less light absorbing properties~~lower lidar ratio values.

2.2.2 Marine

Marine particles are produced at the sea surface and dominate the shallow boundary layer over the oceans (e.g., O'Dowd and de Leeuw, 2007). Specifically, the sea-salt particles feature a predominant coarse mode, however, they are spherical in humid conditions and weakly absorbing in contrast to the dust particles. Therefore, they yield low particle lidar ratio values, are almost non depolarizing and exhibit low Ångström exponent values (e.g., Burton et al., 2014; Dawson et al., 2015). This aerosol type is mainly identifiable by the low particle lidar ratio, i.e., 15–25 sr at 532 nm (Burton et al., 2012). As marine aerosol layers manifest themselves over water bodies, either only stations located at the shorelines and under specific meteorological conditions or shipborne measurements can observe pure maritime particles. Consequently, the observation of pure maritime particles is rare within EARLINET and, generally, when these particles are observed their characteristics are far from pristine (Preißler et al., 2013; Papagiannopoulos et al., 2016a). However, mixtures with important contribution of marine particles can be observed in the Mediterranean basin (Papagiannopoulos et al., 2016a).Thus, we consider pure marine and marine dominated layers as one single category denoted as mixed marine.

2.2.3 Mineral dust and dust mixtures

Mineral dust is produced in arid and semi arid regions of the world, and has a profound contribution to the total natural aerosol loading (Ginoux et al., 2001). The optical properties are considerably different from the other types, thus making them easy to identify. The irregular shape and the large size ($< 50\mu m$; Mahowald et al., 2014) lead to a significant high depolarization of the backscattered radiation (e.g., $\delta^{532}_{aer} = 0.34 \pm 0.02$ for Saharan dust over Germany, Wiegner et al., 2011), and to medium lidar ratio values (e.g., $S^{532}_{aer} = 55 \pm 10\,\mathrm{sr}$, Tesche et al., 2013; Mona et al., 2014). They are spectrally neutral to backscatter and extinction, and thus produce low Ångström exponent values (Wiegner et al., 2011). Therefore, the particle lidar ratio, particle linear depolarization ratio, and the Ångström exponent are excellent physical parameters to characterize mineral dust and to distinguish it from other aerosol types. However, it needs to be taken into account that the dust optical properties depend on the source region and the transport pattern (Valenzuela et al., 2014), which is a source of variability ~~mainly~~ detected in the lidar ratio (e.g., Schuster et al., 2012; Nisantzi et al., 2015). Recently, Mamouri et al. (2013) showed that dust originating from the Arabian desert produced significantly lower lidar ratio values (34–39 sr at 532 nm) than respective values (50–60 sr at 532 nm) from western Saharan dust particles. An overview on the dust characterization using lidar measurements can be found in Mona et al. (2012b).

Dust can be transported over continental scales. In particular, Saharan dust outbreaks to Europe and across the Atlantic Ocean have been deeply investigated. The European continent is regularly influenced by advected Saharan particles as has been discussed by e.g., Ansmann et al. (2003); Guerrero-Rascado et al. (2008, 2009); Papayannis et al. (2008); Müller et al. (2009); Córdoba-Jabonero et al. (2011); Preißler et al. (2011); Valenzuela et al. (2012); Papayannis et al. (2014); Binietoglou et al. (2015); Bravo-Aranda et al. (2015); Granados-Muñoz et al. (2016a). The study of Papayannis et al. (2008) indicated a large variability of the measured lidar ratio and Ångström exponent values among the different sites, suggesting mixing at different levels. Additionally, the mixture processes also produce large variability of intensive properties as measured at the same site (e.g., Mona et al., 2006, 2014). As a consequence of the complex structure of the observed aerosols over Europe and the effects of transport and mixing on the properties of these particles we consider the use of three dust groups: pure dust, mixed dust and polluted dust. The pure dust group refers to particles for which the mixing with other aerosol types is negligible. Mixed dust refers to dust dominated layers mixed with marine particles. This leads to less depolarizing, and less absorbing particles with respect to pure dust particles. ~~Papagiannopoulos et al. (2016a) found this mixture to be important in the Mediterranean region~~ Several studies (Burton et al., 2012; Kim et al., 2013; Rogers et al., 2014; Papagiannopoulos et al., 2016a) have indicated that this mixture is important and suggested its inclusion in the CALIPSO retrieval scheme for improving the accuracy of aerosol backscatter and extinction coefficient profiles. Finally, the polluted dust category consists of dust dominated mixtures with smoke and/or continental, which produce lower depolarization, higher lidar ratios and enhanced Ångström exponent values owing to the presence of small, spherical particles (Groß et al., 2011; Burton et al., 2012; Tesche et al., 2013; Bravo-Aranda et al., 2015).

2.2.4 Smoke

Biomass burning is a major global source of atmospheric aerosols. Generally, smoke particles are relatively small ~~, spherical, and highly absorbing~~ and spherical that produce low depolarization, high Ångström exponents, and large lidar ratios (Amiridis et al., 2009; Baars et al., 2012; Giannakaki et al., 2016). The optical properties of smoke particles may vary due to the vegetation type of the emitting source, the combustion type (smouldering or flaming fires), and atmospheric conditions (e.g., Balis et al., 2003). Furthermore, the particles are susceptible to changes during their lifetime in the atmosphere (Nicolae et al., 2013). Several EARLINET-based studies have focused on observations and characterization of smoke plumes (e.g., Müller et al., 2005; Papayannis et al., 2008; Ansmann et al., 2009; Tesche et al., 2011; Alados-Arboledas et al., 2011; Nepomuceno Pereira et al., 2014; Ancellet et al., 2016; Ortiz-Amezcua et al., 2017), demonstrating that it is a frequently encountered aerosol type over Europe. In particular, biomass burning aerosol originating from forest fires in Canada and Siberia is regularly observed between May and October (Amiridis et al., 2009; Sicard et al., 2012a; Ortiz-Amezcua et al., 2017). However, the similarities of the physical characteristics of smoke particles and continental particles result in similar optical properties, making these types difficult to distinguish. In this work, biomass burning particles are treated as a single category called smoke.

2.2.5 Volcanic ash

Volcanoes are another important source of atmospheric aerosols. Volcanic eruptions eject great amounts of material in the atmosphere (tephra), while the fraction smaller than $2\,\mathrm{mm}$ is labeled as volcanic ash. Most of these aerosols will settle only a few tens of kilometres away from the volcano but smaller particles can travel thousands of kilometres and affect wider areas (Mattis et al., 2010; Sawamura et al., 2012; Sicard et al., 2012b; Navas-Guzmán et al., 2013; Kokkalis et al., 2013; Pappalardo et al., 2013). The optical properties of volcanic ash aerosols is generally similar to the one of desert dust, as was shown by Ansmann et al. (2010) and Wiegner et al. (2012) for fresh ash with particle linear depolarization ratios reaching 0.37 and lidar

~~ratios~~ ratio at 532 sr of 50–65 sr. Aged volcanic particles as observed by Papayannis et al. (2012) indicate less non-sphericity with depolarization ratio values of 0.1–0.25 and lidar ratios for 355 nm within the range 55–67 sr and for 532 nm 76–89 sr. More details can be found in Mona and Marenco (2016) where the authors give a summary of how the intensive optical properties vary as a function of time. Furthermore, volcanic eruptions inject sulfur dioxide into the atmosphere thus leading to sulfate particles. ~~Pappalardo et al. (2004)~~Pappalardo et al. (2004a) and Wang et al. (2008) reported lidar ~~ratios~~ ratio at 355 sr of 50–60 sr and backscatter-related Ångström exponent (355,532) of 2.7, signature of sulfate particles originating from Mount Etna, Italy. Moreover, CALIPSO measurements indicated low particle depolarization ratio for sulfate-rich volcanic clouds (Prata et al., 2017). Consequently, the difference in the optical properties make lidar a powerful tool for volcano monitoring. However, in this study sulfate particles and aged volcanic particles are not considered. The aerosol type relevant to the airborne ash refers to fresh ash and is denoted as volcanic ash.

As an additional consideration, the defined aerosol types presented in Sect. 2.2 may not be representative of the entire aerosol load and, furthermore, apart from the dust mixtures they do not consider other aerosol mixtures. For example, this aspect can be observed in the definition of the volcanic category where the particles have different characteristics depending the transport pattern. The particles near the source have optical properties similar to desert dust whereas long-range transported volcanic plumes have the altered properties due to the sedimentation of the coarser particles. Therefore, it is important to further include a more exhaustive aerosol class analysis.

3 Automatic aerosol type classification

3.1 Methodology

We developed an automated typing method, based on the work of Burton et al. (2012), but modified it in order to be compatible with the database of EARLINET. ~~We used EARLINET data from a $3\beta + 2\alpha$ setup, hence, data retrieved during nighttime operation. Doing this, we can estimate the maximum number of intensive parameters relevant for the classification and, furthermore, potentially apply the classification to historical EARLINET data. Besides, we explore the inclusion of depolarization ratio observations in the automatic algorithm. The selection of the aerosol types is based on the major aerosol components found over EARLINET sites and it is, also, examined the possibility of combining aerosol types to obtain better results.~~

~~The method can be separated into two important steps~~Two major steps are identified in the method proposed: the training (Sect. 3.2), and the testing (Sect. 4.1) phase. The first step consists of the following procedures. As described in Sect. 3.2.1, well characterized aerosol ~~samples~~ layers are manually separated into classes based on their physical characteristics~~and this presents the~~; the set of classesconstitutes the reference dataset. This procedure involves the determination of each observed aerosol layer location and the estimation of mean layer intensive optical properties. Based on this analysis, the classifying parameters that provide the ~~adequate~~ required information for a better discrimination of the aerosol type are selected (Sect. 3.2.2). Next, in order to estimate how accurately a predictive model will perform, the reference dataset is split into training and validation datasets, and the application of the classifier is evaluated (Sect. 3.2.3). Sect. 3.2.4 describes the inference of characteristic depolarization values in the algorithm with the intention to increase the prediction of the model.

For the second step, ~~unclassified EARLINET data (the testing dataset) is categorized using the reference dataset, that is, aerosol layers are identified and their mean layer intensive properties are entered in the automatic classification scheme. Besides, these data have been classified following the method shown in Sect. 2.1 and, hence, compared against the output~~ already pre-classified EARLINET data are used to assess the performance of the automatic typing procedure. Figure 4 illustrates the sequence of the proposed methodology starting from the setting of the training dataset, up to the assessment of the learning success during the testing phase.

Distance-based classification methods aim to assign an observation to a particular class based on the distance of the observation from each class center. In general, the Mahalanobis distance between an observation $x = (x_1, ..., x_p)^t$ and the mean class $\bar{x} = (\bar{x}_1, ..., \bar{x}_p)^t$ in the p-dimensional space $\mathbb{R}^p$ is defined as

$$\boldsymbol{D}_M(x, \bar{x}) = \sqrt{(\boldsymbol{x} - \bar{\boldsymbol{x}})^T \boldsymbol{S}^{-1} (\boldsymbol{x} - \bar{\boldsymbol{x}})}. \tag{3}$$

Where $\boldsymbol{S}$ is the class covariance matrix. The surfaces identified by the equation $D_M = const.$ are ellipsoids that are centered around the mean $\bar{x}$. The main characteristic of the multivariate Mahalanobis distance is that it accounts for the variance of each variable and the covariance between variables. By contrast, the Euclidean distance treats all the variables in the same way and the constant distance surfaces from a fixed point are represented by a sphere.

The Mahalanobis distance of an observation from an aerosol class is estimated, and ~~it~~ is assigned to the aerosol class for which the distance is minimum. ~~The minimum accepted distance is set to a threshold assuming that the statistical distance belongs to a chi-squared distribution. The selected threshold distance represents the 99.9 % cumulative probability contour of the class distribution and varies with the degrees of freedom (i. e., the classifying parameters). In this work, the minimum distance~~ Two screening criteria are applied to the minimum distance following the procedure of Burton et al. (2012). The methodology uses 3 and 4 classifying parameters and the minimum accepted distance for a measurement to be labelled is 4 and ~~corresponds to a 3-dimensional space, considering 3 classifying parameters~~4.3 respectively. Moreover, the ~~estimated Mahalanobis distances for each class can be assigned to a probability. These probabilities are normalized using the sum of each class probability and a measurement point is labelled if the relative probability is greater~~ normalized probability of the aerosol class needs to be higher than 50 %. Otherwise, the type assignment is difficult as the measurement can be equidistant from 2 or more aerosol type classes, and possibly indicate the mixing of these aerosol types. ~~In the future, this information could be used to label complicated aerosol scenes with different level of aerosol mixing.~~

3.2 Training phase

3.2.1 Dataset

In supervised learning techniques, the reference dataset is crucial to the overall predictive performance of the algorithm. Therefore, it is fundamental to use well-characterized EARLINET profiles. Namely, EARLINET aerosol classified layers from Pappalardo et al. (2013); Papagiannopoulos et al. (2016a); Schwarz (2016) were used and will be presented below.

EARLINET observations from 2008 to 2010 were analyzed and the aerosol types were determined with respect to the source origin following a similar approach to Sect. 2.1 (Schwarz, 2016) and present the backbone of the reference dataset. Table 1 lists the classified aerosol types of the above study (644 individual aerosol layers) with respect to the aerosol types presented in Sect. 2.2; however, all these aerosol layers cannot be used given the need for the maximum optical properties available (column "only from $3\beta + 2\alpha$"). The mixtures category includes all the mixtures of two or more aerosol species without containing polluted dust and mixed dust categories that are reported individually.

As discussed above, the requirement for $3\beta + 2\alpha$ lidar configuration pinpoints the low occurrence (see Table 1) of some aerosol types such as the clean continental, polluted dust, and dust. Furthermore, marine aerosol was not reported in the study and the volcanic layers do not reflect the volcanic ash characteristics described in Sect. 2.2. Conversely, the latter were volcanic layers found in the stratosphere and, thus, different from the fresh ash that we consider. In order that the aerosol classes include all the major aerosol components, the aforementioned aerosol types need to be enhanced with other observations. Therefore, we implemented EARLINET network-wide typing results already published in literature (Pappalardo et al., 2013; Papagiannopoulos et al., 2016a) for a total of 69 ~~samples~~layers, as the reference dataset. Note that calibrated particle linear depolarization ratio profiles are not available in the selected dataset.

The type-dependent mean properties are reported in Table 2 and coincide with the typical values as of those in Sect. 2.2. However, aerosol classification is based on an interpretative analysis of the retrieved optical properties and the model simulations, and, it is a qualitative method of type assignment. Thus, there is inherent possibility of error in the determination of the true aerosol type. This error, if made, propagates into the automatic algorithm and the predicted aerosol class might deviate from the "truth" aerosol class. Specifically, dust and volcanic types present the same characteristics with Ångström ~~exponent~~ exponents as low as 0, although dust lidar ratio is 58 ± 12 and 55 ± 7 sr, for 355 nm and 532 nm respectively, and is higher than the volcanic lidar ratio (S_{aer}^{355}=50 ± 11 and S_{aer}^{532}=48 ± 13 sr). The Ångström ~~exponent~~ exponents (i.e., $\kappa_\beta(355, 1064)$, $\kappa_\beta(532, 1064)$, $\kappa_\beta(355, 532)$, and $\kappa_\alpha(355, 532)$) for mixed dust ~~is~~ are between 0.4–0.7 and lidar ratio values are below 50 sr, whereas for polluted dust the Ångström ~~exponent lies~~ exponents lie within 0.6–1.0 and lidar ratio values for 355 nm and 532 nm are 54 ± 8 and 64 ± 9 sr. This behavior reflects the mixing of dust with pollution/smoke that tends to decrease the size of the aerosol mixture and increase its absorbing capacity. Polluted continental and smoke reveal the same size characteristics with mean Ångström exponent from all the available variables around $\sim$1.4 and $\sim$1.3 respectively. The smoke mean lidar ratio values present the higher ones among the aerosol types – i.e., 81 ± 16 sr and 78 ± 11 sr for 355 nm and 532 nm respectively – and the polluted continental values succeed with 69 ± 12 sr and 63 ± 13 sr for 355 nm and 532 nm respectively. For clean continental the Ångström ~~exponent is~~ exponents (i.e., $\kappa_\beta(355, 1064)$, $\kappa_\beta(532, 1064)$, $\kappa_\beta(355, 532)$, and $\kappa_\alpha(355, 532)$) are between

1.0–1.7 and lidar ratios, for 355 nm and 532 nm, are 50 ± 8 sr and 41 ± 6 sr. This characteristic separates clean continental from polluted continental as the particles ~~are less absorptive~~yield lower lidar ratio values. Finally, mixed marine particles are found to be relatively small in size with Ångström exponents (i.e., $\kappa_\beta(355, 1064)$, $\kappa_\beta(532, 1064)$, $\kappa_\beta(355, 532)$, and $\kappa_\alpha(355, 532)$) in the range 0.8–1.0 and, thus, overlap with other aerosol types. The characteristic parameter that defines the mixed marine category is the lidar ratio, the values are found the smallest – S_{aer}^{532}=24 ± 8 sr – among the aerosol types.

In the proposed method, the aerosol layers are classified in terms of the aerosol types described in Sect. 2.2. As a starting point for this study, we use 8 aerosol classes: clean continental (CC), polluted continental (PC), pure dust (D), mixed dust (MD=Dust+Marine), polluted dust (PD=Dust+Smoke and/ or Dust+Polluted Continental), mixed marine (MM), smoke (S), and volcanic (V). However, some of these 8 classes overlap consistently in the feature space. As a consequence, we exploited the combined use of overlapping aerosol types. Therefore, we merged the types that tend to reflect the same aerosol characteristics, and hence, we evaluate the corresponding effects on the prediction rate of the algorithm. Two pathways were followed, first, the smoke and the polluted continental categories were grouped into the more generic type of small ~~, absorbing particles,~~ with high lidar ratio values, and, second, all the dust-like aerosol types were merged. The different grouping categories are summarized in Table 3.

3.2.2 Classifying parameters selection

Next, we performed a sensitivity analysis to identify which classifying properties provide the adequate information to better predict the correct aerosol class. We used three aerosol intensive properties due to the lack of particle linear depolarization ratio profiles to evaluate the strength of the selected ~~classifiers~~ classifier to discriminate among the predefined classes. Two statistical parameters are used: the total and the partial Wilks' lambda (Λ; Wilks, 1963) that are widely used e.g., Burton et al. (2012) and Russell et al. (2014). The total Λ statistic shows the tendency of the above set of pre-specified classes (or any subset of it) to separate. The partial Λ is calculated for each of the intensive properties separately and indicates the discriminatory power of the used intensive property. For both parameters, values range from 0–1. Values near zero show high discriminatory power while values near one show low discriminatory power.

The lowest total Λ was found to be 0.033 for the set $\kappa_\beta(355, 1064)$, S_{aer}^{532}, and $S_{aer}^{532}/S_{aer}^{355}$; whereas the partial Λ is 0.51 for $S_{aer}^{532}/S_{aer}^{355}$, 0.17 for κ_β, and 0.30 for S_{aer}^{532}. For this dataset, the low Λ value for κ_β indicates that this variable has the most weight in the classification. The decision for the selected parameters stems solely to the lowest arithmetic value of the total Λ. Therefore, for the other pairs of parameters the total Λ is equally low, ~0.05, which indicates that, also, a $2\beta + 2\alpha$ lidar setup could be equally used when the algorithm is trained with $\kappa_\beta(355, 532)$. With reference to the lidar ratio, the S_{aer}^{532} and S_{aer}^{355} can be used interchangeably due to the almost equal total Λ (i.e., 0.034).

For the rest aerosol groups reported in Table 2 the total and partial (for κ_β) Λ are 0.036 and 0.18 respectively (7^a classes), 0.041 and 0.18 (7^b classes), 0.044 and 0.18 (6 classes), 0.057 and 0.20 (5 classes), and 0.070 and 0.21 (4 classes). The Λ shows good discriminatory power for each of the grouping classes, although there is a slight increase in the values as the number of classes is reduced. This behavior can be ascribed to the high variance of the combined aerosol types which makes the classification less selective.

Figure 5 shows the characteristics of the reference dataset in terms of the ~~lidar ratio and Ångström exponent~~ S_{aer}^{532} and $\kappa_{\beta}(355, 1064)$ for the 8 and 4 aerosol classes that represent the maximum and minimum used aerosol groupings. The coloring corresponds to the various classes and the crosshairs indicate the standard deviation of each of the aerosol layers. The 90 %-confidence ellipses are calculated using the eigenvalues and eigenvectors of the covariance matrix and define the region that contains the 90 % of all the points that can be drawn from the underlying normal class distribution. The various aerosol classes tend to populate specific areas of the graph whereas the overlap of the neighbouring classes is significant, although the classes are better pinpointed as long as we merge classes with similar characteristics. The latter does not reflect the obtained values of the statistical parameters (total Λ increased from 0.033 to 0.070 for 8 classes and 4 classes respectively), however, and as explained above, the reference dataset very well delineates the aerosol types and by combining the neighbouring types the variance increases.

3.2.3 Validation of the classifier

In order to evaluate the predictive accuracy of the automatic method it is needed to split the initial reference dataset into a training and a validation dataset. Like this, we use the training dataset to calculate the classification functions and then submit each observation in the validation dataset to the classification function obtained from the training dataset. For this study, we make use of the leave-one-out cross validation (LOOCV) procedure, also referred to as holdout procedure or simply cross validation, which is a degenerate case of the k-fold cross validation, where k is chosen as the total number of samples (Rencher, 2002). The choice of the procedure, even though computational expensive, is used when datasets are sparse and trains the algorithm with as many observations as possible. Each measurement separately is removed from the training dataset in order to compute the classification rule, and this rule is used to classify the removed observation. The error rate is estimated as a ~~percentage of errors made over the whole set of observations used for testing.~~ percentage of all incorrect predictions divided by the total number of the reference dataset, and is equivalent to 1 minus accuracy. Values near zero show high predictive performance while values near one show low predictive performance. For the classification options of the Table 3, the error rate, expectedly, decreases with decreasing number of aerosol classes (39 % for 8 classes, 36 % for 7^a classes, 30 % for 7^b classes, 28 % for 6 classes, 19 % for 5 classes, and 10 % for 4 classes). It should be mentioned that the typing in multiple classes and typing accuracy are two conflicting aspects. The choice of 8 aerosol classes appears to be sufficient to describe the major aerosol components, however ostentatious for a $3\beta + 2\alpha$ lidar configuration. 4 classes, on the other hand, provide a coarse aerosol characterization and the prediction accuracy of the algorithm is expected to increase.

3.2.4 Algorithm training including particle depolarization ratio

Several studies have shown the unique information provided by depolarization measurements (e.g., Liu et al., 2008; Tesche et al., 2013; Burton et al., 2014), thus making this intensive property a robust means to discriminate the various aerosol types. Valuable typing information can also be obtained by the color ratio of the particle depolarization ratios when more depolarization channels exist ~~(Burton et al., 2014)~~(Burton et al., 2015). As already stated in Sect. 2, the majority of the stations perform depolarization measurements, and profiles are routinely delivered by SCC. However, the reference dataset does not contain depolarization information because it has been released before the assessment of the quality assurance procedures within EARLINET. Therefore, a method applicable to EARLINET data collected since 2000 is proposed in this work. We investigate the effect of adding depolarization information to the described method as the next releases of EARLINET dataset will contain quality assured particle depolarization profiles and can be used for more accurate aerosol typing. To complement the reference dataset in this context, we used general literature values for particle linear depolarization ratio at 532 nm (Table 4) in order to train the algorithm. For the clean continental type, the values ingested in the algorithm are retrieved from Burton et al. (2013) and refer to the polluted marine category. The decision for this inconsistency stems to the shortage of clean continental particle depolarization values in literature, however the reported values coincide with the type characteristics described in Sect. 2.2 and the values used in the CALIPSO typing scheme (Omar et al., 2009) ~~.~~

In this case, the particle linear depolarization ratio ~~replaced the worst performing classifier – i.e., the $S^{532}_{aer}/S^{355}_{aer}$. Hence, the three classifying parametersare the δ^{532}_{aer}, S^{532}_{aer}, and $\kappa_\beta(355,1064)$~~was added to the classifying parameters. Values within the aerosol type range were randomly assigned to each sample and the Λ distribution was calculated. Total Λ is ~~0.008~~0.004. The value of partial Λ for ~~κ_β~~$\kappa_\beta(355,1064)$, S^{532}_{aer}, $S^{532}_{aer}/S^{355}_{aer}$ and δ^{532}_{aer} are ~~0.69, 0.36,~~ 0.55, 0.34, 0.52 and 0.12 respectively. The values found for the partial Λ confirm the δ^{532}_{aer} as the most important ~~classifier~~ classifying parameter for the considered dataset. For the rest aerosol groups the total and partial (for depolarization ratio) Λ are ~~0.009 and 0.13~~ 0.005 and 0.14 respectively (7^a classes), ~~0.010~~ 0.005 and 0.12 (7^b classes), ~~0.012 and 0.13~~ 0.006 and 0.14 (6 classes), ~~0.072 and 0.67~~ 0.040 and 0.68 (5 classes), and ~~0.090~~ 0.050 and 0.69 (4 classes).

For the sake of completeness, the LOOCV method was also performed and the error rate was calculated. Figure 6 presents comparatively the training of the algorithm when depolarization information is available and when not in terms of the total, and partial Λ and the error rate of the LOOCV method. The figure highlights the strength of polarization sensitive observations, while for the 5 and 4 classes (Fig. 6b) the particle linear depolarization ratio becomes less important (in this case the highest weight in the classification corresponds to the lidar ratio at 532 nm) due to the fact that only one dust type represents volcanic and other dust mixtures. Figure 7 presents cumulative bar plots with the median (black dots), the 25–75 percentile (box), the 5–95 percentile (whiskers) for all 4 classifying parameters. The figure highlights the discriminatory power of δ^{532}_{aer}, $\kappa_\beta(355,1064)$, and S^{532}_{aer}, whereas the $S^{532}_{aer}/S^{355}_{aer}$ performs the worst. Furthermore, the figure depicts the discriminatory power of the classifying parameter among the dust-like aerosol classes, however, the particle depolarization ratio seems to have no power to separate the non-dust classes, as discussed above.

4 Results

4.1 Testing phase

As a next step, an assessment of the predictive performance of the pre-trained algorithm is made by using a testing dataset. For this, EARLINET data collected during the ACTRIS Summer 2012 intensive measurements (Sicard et al., 2015; Granados-Muñoz et al., 2016b) were chosen to test the automatic typing algorithm. The measurements took place in the period of 8 June–17 July 2012 and were dedicated to Saharan dust studies and also featured two field campaigns such as PEGASOS (Pan

European Gas Aerosols Interaction Study) and Charmex (Chemistry Aerosol Mediterranean Experiment). During that period, 157 measurements were performed, out of which 42 measurements delivered 3 backscatter and 2 extinction coefficient profiles. The description of aerosol type distribution over Europe during the campaign was obtained following the procedure shown in Sect. 2.1 (Papagiannopoulos et al., 2016b). The testing dataset comprises of 47 ~~samples~~layers, 21 of which yield depolarization ratio values. Table 5 provides the mean values of the intensive parameters for each available category in accordance with Table 2.

4.2 Application of the methodology to EARLINET data - case studies

To showcase the steps of the automatic classification, we apply it to two selected cases for the 8 classes and for the classifying parameters: $S^{532}_{aer}/S^{355}_{aer}$, S^{532}_{aer}, and $\kappa_\beta(355, 1064)$. For the case in Sect. 2.1, the automatic algorithm labelled the aerosol layer as ~~pure dust, $D_M = 1.0$~~ dust, $D_M = 1.2$ and the normalized probability ~~62~~55 %.This coincides with our findings and highlights the strength of the classification, albeit this example corresponds to a pure aerosol layer with no level of mixing with other aerosol types.

The second case refers to a more complicated aerosol scene. The Athens EARLINET station (Fig. 8) on 22 May, 2014 observed an aerosol layer mostly in the height range between 1.5 and 3 km (Papayannis et al., 2016). Within this layer the mean value of ~~backscatter (extinction) related~~ backscatter-related Ångström exponent (355,1064) is ~~0.9 ± 0.1 (1.0 ± 0.4)~~ 0.9 ± 0.1. The lidar ratio presents mean values in the layer 40 ± 7 sr and 39 ± 6 sr at 355 nm and 532 nm, respectively. The color ratio of the lidar ratios shows a wavelength independent layer with values 1.1 ± 0.2. The retrieved error corresponds to the standard deviation of the retrieved quantity calculated within the layer.

In the following, a 6-day FLEXPART backward trajectory indicates the pattern of the origin of airmasses. Figure 9 shows the total column sensitivity of the particles found over the station in between 1.5–3 km, it highlights the motion of the particles in a north-easterly direction towards the Aral Sea and Kazakhstan. This area is an active dust source due to the extreme desiccation of the lake (Ginoux et al., 2012). Therefore, the path of the air masses arriving over Athens suggests a mixture of dust ~~, marine~~ and biomass burning particles, originating from the arid areas of the Aral Sea, as well as the agricultural fires in former Soviet Union countries ~~, enriched with marine particles during their overpass over the Black Sea~~ (Papayannis et al., 2016). The automatic algorithm classified the layer as mixed dust, $D_M = 2.5$ and normalized probability 32 %, and the second closest class was clean continental, $D_M = 3$ and normalized probability 23 %. Although the class with the minimum estimated distance agrees with our investigation, the inferred type will not be taken into account. The very low probability indicates that more than one distance is beyond the accepted threshold, therefore the classes are almost equidistant. This demonstrates that the manual typing procedure can better type the aerosol layer, but also that adopted fixed thresholds are conservative, i.e., type assignment is not possible for ambiguous scenes.

4.3 Comparing the automatic classification with manual analyzed data

The performance of the algorithm with respect to the testing dataset is presented. For each of the grouping classes, as those listed in Table 3, the confusion matrices have been calculated (not shown) and the accuracy of the model is presented, here, alongside with the recall (R) and precision (P). The confusion matrix describes the performance of the classifier on a testing dataset for which the typing is already known. ~~Accuracy shows how many times the predicted aerosol type agrees with the true aerosol type and is the ratio of the correctly classified instances (true positives) to the total number of instances. Recall of each~~ Recall of an aerosol group is defined as the number of ~~true positives (T_p)~~ correctly predicted cases over the number of ~~true positives~~ correctly plus the number of ~~false negatives (F_n). The false negatives represent the instances of cases when we predicted a different aerosol type than the true one.~~

$$R = \frac{T_p}{T_p + F_n}$$

~~Precision of each~~ incorrectly predicted cases. Recall can be thought as the model's ability to predict the specific aerosol class. Precision of an aerosol group is defined as the number of ~~true positives (T_P)~~ correctly predicted cases over the number of ~~true positives~~ correctly predicted cases plus the number of ~~false positives (F_p). The false positives represent the number of instances when we predicted a specific aerosol type although different from the true one.~~ incorrectly predicted cases that belong to this aerosol class. In other words, given the prediction of a specific class, what is the probability of being correct?

$$P = \frac{T_p}{T_p + F_p}$$

~~For an ideal classification procedure both $F_p = F_n = 0$ and so $R = P = 1$. For real classification systems false positive and/or false negatives may occur and consequently values of R and P between 0 and 1 are possible. The more the value of R and P is far from 1 the less the classification system behaves like an ideal one. In particular, a value of R (P) close to 1 indicates the generation of negligible number of false negatives (positives) during the classification procedure.~~

In Fig. 10, the bar plot shows comparatively the predictive accuracy of the algorithm when compared to manual analyzed data for the different aerosol classes in both the cases in which the depolarization information is available (in orange) or not (in brown). Without depolarization ratio information, the accuracy of the model increases with decreasing number of classes. The lowest value was obtained for 8 classes (59 %) and the highest for 4 classes (90 %). With depolarization ratio information, the accuracy for 8 classes equals to ~~69~~79 % and exceeds the 80 % for the rest aerosol classes. When comparing the accuracy of the model with and without depolarization ratio, it appears to be significantly higher until 6 classes where, further, the discrepancy diminishes (~~$\sim$~~$<$10 %) and becomes smaller for 4 classes. In general, it becomes evident that the particle linear depolarization ratio increases the ability for predicting correctly the aerosol type. Given the high accuracy, a $3\beta + 2\alpha$ configuration showed that 6 aerosol classes, as well as 5 and 4, can provide a robust classification. Instead, the training of the classification with depolarization measurements enhances the predictability strength and can provide finer aerosol classification (for 8 classes, accuracy $\sim$70 %).

Table 6 summarizes the results when using as classifying parameters: $S^{532}_{aer}/S^{355}_{aer}$, S^{532}_{aer}, and $\kappa_\beta(355, 1064)$, with respect to recall and precision and offer a better insight in the performance of each aerosol type. Next to the number of classes, between parentheses, the number of aerosol layers that passed the screening criteria as those described in Sect. 3.1 is provided. It is, thus, worth noting that the numbers increase when the aerosol types are combined. The mixed marine and clean continental aerosol types yield high recall and precision (values>80 %) throughout the different aerosol classes, highlighting the ability of the classifier to correctly label them. The aerosol types that performed worse are the smoke and polluted continental aerosol types due to the similarities in the intensive optical properties. However, when combining them into a single aerosol class (see 7^b, 6, and 4 classes) precision and recall increase significantly. Given the noticeable signature of dust particles precision is high, whereas the recall is 30 % and this can be assigned to the lack of depolarization measurements. Similarly, recall increases as soon as volcanic, mixed and polluted dust are included in the same all-dust category (see 7^a, 6, 5, and 4 classes). Note that mixed dust and polluted dust aerosol types are not reported in the tables due to the fact that they are not present in Table 5 and these parameters cannot be evaluated. Note that mixed dust and polluted dust aerosol types are not reported in the tables due to the fact that they are not present in Table 5 and these parameters cannot be evaluated. The frequency of detection for MD (PD) is 18 % (4 %) for 8 classes, 15 % (3 %) for 7^b classes, 17 % (3 %) for 7^a classes, and 15 % (3 %) when 3 classifying parameters are used. The algorithm predicted as MD only dust cases with S^{532}_{aer} around 45 sr and $S^{532}_{aer}/S^{355}_{aer}$ over 1. The PD case refers to a PC case with $\kappa_\beta(355, 1064)$ lower than 1. The frequency of detection for MD is 0 % for all classes when depolarization ratio is added. The frequency for PD is 17 % for 8 classes, 13 % for 7^b classes, 17 % for 7^a classes, and 13 % for 6 classes. The wrongly classified cases have depolarization ratio around 20 %.

Table 7 is similar to Table 6 and reports the recall and precision when depolarization information is available. Clean continental aerosol, again, yields high recall and precision for all the different aerosol groups. Polluted continental performed the worse and, expectedly, showed the same behavior as before when compared with smoke into a single type. Alternately, dust is precisely identified for all the aerosol classes. This result indicates that depolarization measurements facilitate the correct dust typing. It is noteworthy, that although the findings are promising the test dataset is limited and do not cover all the aerosol classes.

5 Summary and Conclusions

The characterization of the vertical aerosol distribution is needed for accurate radiative-transfer modeling. Automatic procedures to classify aerosols objectively and within near-real-time scales are employed. An automatic classification procedure based only on EARLINET data was presented. Here, we modified an automatic algorithm to satisfy the network's requirements and needs. A Wilks' lambda analysis was performed on EARLINET data and the three best performing classifying parameters were: the lidar ratio at 532 nm; the color ratio of the lidar ratios at 355 nm and 532 nm and the backscatter-related 355-to-1064- nm Ångström exponent. Nevertheless, the other intensive parameters using the available wavelengths can be equally used as the analysis showed similar values. Furthermore, the number of aerosol classes has been investigated for a maximum of 8 and minimum 4. Prior to evaluating the performance of the algorithm, the leave-one-out cross validation procedure was performed on the reference dataset and the error rate decreased monotonously from 39 % to 10 % with decreasing number of aerosol classes. The prediction of the automatic classification showed positive results when compared against already classified

EARLINET data. In particular, the positive learning success for 8 (59 %), 7 (69 % for 7^a and 7^b classes), 6 (76 %), 5 (76 %) and 4 (90 %) classes indicates that the fewer aerosol classes (6, 5, and 4 classes) provide a confident but, nonetheless, a coarser classification. To be more precise, the high accuracy (76 %) coupled with the low error rate of the cross-validation (28 %) for 6 classes offers a good starting point for a classification with a $3\beta + 2\alpha$ lidar configuration.

Besides, the training of the algorithm with literature depolarization ratio values decreased the error rate of the leave-one-out cross validation from ~~30~~24 % ~~to 7~~(8 classes) to 4 % ~~and increased the predictive accuracy. For 8 classes~~(4 classes). Furthermore, the predictive accuracy ~~reached 69 % while for the rest it remained well above~~ increased and remained for all the aerosol classes around 80 % (for 8 classes: 79 %, for 7^a: ~~86~~81 %, for 7^b: 83 %, for 6: ~~93~~81 %, for 5: ~~86~~85 %, and for 4: ~~93~~80 %). Therefore, this finding suggests that the algorithm in this case can be used for finer aerosol classification and also delineates the discriminatory power of depolarization ratio. Specifically, 7 aerosol classes (either D+V, MD, PD, CC, MM, PC, S or D, V, MD, PD, CC, MM, PC+S) seem to be adequate to provide reasonable typing results. However, the obtained results refer to a small testing dataset that consists of pure aerosol types and underestimates the aerosol mixtures of the classification.

The presented automatic algorithm is based only on EARLINET data and is set to accommodate EARLINET measurements covering as much of its measurement record as possible. Specifically, Raman lidar systems with $3\beta + 2\alpha$ and $2\beta + 2\alpha$ configuration with and without particle depolarization ratio can be used for the aerosol classification. The manageability of the algorithm regarding the reference dataset, the number of the aerosol classes and the classifying parameters make the method easily adaptable and handled by individual users. The training dataset can be easily enlarged with high quality typing data coming from a multitude of EARLINET stations and a longer time record. Moreover, new classifying parameters, particle linear depolarization ratio at more wavelengths and aerosol extinction coefficient in the infrared, can be easily added as the observing capacity increases.

The use of the method network-wide will homogenize and standardize the aerosol typing towards a new EARLINET product. The implementation of the method into the SCC will create a complete automatic lidar analysis, that is, from the retrieval of optical properties to aerosol classification. Further, an intercomparison of the developed method against methods which make use also of aerosol optical property modeling could improve from one side the optimization of aerosol property models and from the other side the tuning of aerosol types and reference dataset. This method, even if developed on the basis of EARLINET and its variable instrumental capability, can be applied to all of the aerosol lidar systems as those part of GALION as well as to future lidar-based satellite missions (e.g., the Earth Cloud Aerosol and Radiation Explorer (EarthCARE) satellite mission). In future, a combination of the few sophisticated EARLINET-type lidars and extended networks of automated single-wavelengths backscatter lidars (as ceilometers, Wiegner et al., 2014) might be beneficial with aerosol typing provided at “anchor stations”, and the spatial extent of the layers can be provided by the continuous observations of the ceilometers. This will also offer a unique data set for evaluation of models.

Acknowledgements. The financial support for EARLINET in the ACTRIS Research Infrastructure Project by the European Union’s Horizon 2020 research and innovation programme under grant agreement no. 654169 in the Seventh Framework Programme (FP7/2007–2013) is gratefully acknowledged. The research leading to these results has received funding from European Union’s Horizon 2020 research and innovation programme under grant agreement no. 602014 (project ECARS (East European Centre for Atmospheric Remote Sensing)) and from European Union's Horizon 2020 research programme for Societal challenges - smart, green and integrated transport under grant agreement no. 723986 (project EUNADICS-AV (European Natural Disaster Coordination and Information System for Aviation)). D. Bortoli acknowledges the European Union through the European Regional Development Fund, included in the COMPETE 2020 (Operational Program Competitiveness and Internationalization) through the ICT project (UID/GEO/04683/2013) with the reference POCI 01-0145-FEDER 007690.

References

Alados-Arboledas, L., Müller, D., Guerrero-Rascado, J. L., Navas-Guzmán, F., Pérez-Ramírez, D., and Olmo, F. J.: Optical and micro-physical properties of fresh biomass burning aerosol retrieved by Raman lidar, and star- and sun-photometry, Geophys. Res. Lett., 38, doi:10.1029/2010GL045999, 2011.

Amiridis, A., Balis, D. S., Giannakaki, E., Stohl, A., Kazadzis, S., Koukouli, M. E., and Zanis, P.: Optical characteristics of biomass burning aerosols over Southeastern Europe determined from UV-Raman lidar measurements, Atm. Chem. and Phys., 9, 2431–2440, doi:10.5194/acp-9-2431-2009, 2009.

Ancellet, G., Pelon, J., Totems, J., Chazette, P., Bazureau, A., Sicard, M., Di Iorio, T., Dulac, F., and Mallet, M.: Long-range transport and mixing of aerosol sources during the 2013 North American biomass burning episode: analysis of multiple lidar observations in the western Mediterranean basin, Atmospheric Chemistry and Physics, 16, 4725–4742, doi:10.5194/acp-16-4725-2016, 2016.

Ansmann, A., Wagner, F., Althausen, D., Müller, D., Herber, A., and Wandinger, U.: European pollution outbreaks during ACE 2: Lofted aerosol plumes observed with Raman lidar at the Portuguese coast, J. Geophys. Res., 106, doi:10.1029/2000JD000091, 2001.

Ansmann, A., Bösenberg, J., Chaikovsky, A., Comeròn, A., Eckhardt, S., Eixmann, R., Freudenthaler, V., Ginoux, P., Komguem, L., Linné, H., Márquez, M. A. L., Matthias, V., Mattis, I., Mitev, V., Müller, D., Music, S., Nickovic, S., Pelon, J., Sauvage, L., Sobolewsky, P., Srivastava, M., Stohl, A., Torres, O., Vaughan, G., Wandinger, U., and Wiegner, M.: Long-range transport of Saharan dust to northern Europe: The 11-16 October 2001 outbreak observed with EARLINET, J. Geophys. Res., 108(D24), doi:10.1029/2003JD003757, 2003.

Ansmann, A., Baars, H., Tesche, M., Müller, D., Althausen, D., Engelmann, R., Pauliquevis, T., and Artaxo, P.: Dust and smoke transport from Africa to South America: Lidar profiling over Cape Verde and the Amazon rainforest, Geophys. Res. Lett., 36, doi:10.1029/2009GL037923, 2009.

Ansmann, A., Tesche, M., Groß, S., Freudenthaler, V., Seifert, P., Hiebsch, A., Schmidt, J., Wandinger, U., Mattis, I., Müller, D., and Wiegner, M.: The 16 April 2010 major volcanic ash plume over central Europe: EARLINET lidar and AERONET photometer observations at Leipzig and Munich, Germany, Geophysical Research Letters, 37, doi:10.1029/2010GL043809, l13810, 2010.

Baars, H., Ansmann, A., Althausen, D., Engelmann, R., Heese, B., Müller, D., Artaxo, P., Paixao, M., Pauliquevis, M., and Souza, R.: Aerosol profiling with lidar in the Amazon Basin during the wet and dry season, J. Geophys. Res, 117, doi:10.1029/2012JD018338, 2012.

Baars, H., Kanitz, T., Engelmann, R., Althausen, D., Heese, B., Komppula, M., Preißler, J., Tesche, M., Ansmann, A., Wandinger, U., Lim, J.-H., Ahn, J. Y., Stachlewska, I. S., Amiridis, V., Marinou, E., Seifert, P., Hofer, J., Skupin, A., Schneider, F., Bohlmann, S., Foth, A., Bley, S., Pfüller, A., Giannakaki, E., Lihavainen, H., Viisanen, Y., Hooda, R. K., Pereira, S. N., Bortoli, D., Wagner, F., Mattis, I., Janicka, L., Markowicz, K. M., Achtert, P., Artaxo, P., Pauliquevis, T., Souza, R. A. F., Sharma, V. P., van Zyl, P. G., Beukes, J. P., Sun, J., Rohwer, E. G., Deng, R., Mamouri, R.-E., and Zamorano, F.: An overview of the first decade of Polly^{NET}: an emerging network of automated Raman-polarization lidars for continuous aerosol profiling, Atmospheric Chemistry and Physics, 16, 5111–5137, doi:10.5194/acp-16-5111-2016, https://www.atmos-chem-phys.net/16/5111/2016/, 2016.

Baklanov, A., Schlünzen, K., Suppan, P., Baldasano, J., Brunner, D., Aksoyoglu, S., Carmichael, G., Douros, J., Flemming, J., Forkel, R., Galmarini, S., Gauss, M., Grell, G., Hirtl, M., Joffre, S., Jorba, O., Kaas, E., Kaasik, M., Kallos, G., Kong, X., Korsholm, U., Kurganskiy, A., Kushta, J., Lohmann, U., Mahura, A., Manders-Groot, A., Maurizi, A., Moussiopoulos, N., Rao, S. T., Savage, N., Seigneur, C., Sokhi, R. S., Solazzo, E., Solomos, S., Sørensen, B., Tsegas, G., Vignati, E., Vogel, B., and Zhang, Y.: Online coupled regional meteorology chemistry models in Europe: current status and prospects, Atmospheric Chemistry and Physics, 14, 317–398, doi:10.5194/acp-14-317-2014, 2014.

Balis, D., Amiridis, V., Nickovic, S., Papayannis, A., and Zerefos, C.: Raman lidar and sunphotometric measurements of aerosol optical properties over Thessaloniki, Greece during a biomass burning episode, Atmos. Environ., 31, 4529–4538, doi:10.1016/S1352-2310(03)00581-8, 2003.

Biele, J., Beyerle, G., and Baumgarten, G.: Polarization lidar: Corrections of instrumental effects, Opt. Express, 7, 427–435, doi:10.1364/OE.7.000427, 2000.

Binietoglou, I., Basart, S., Alados-Arboledas, L., Amiridis, V., Argyrouli, A., Baars, H., Baldasano, J. M., Balis, D., Belegante, L., Bravo-Aranda, J. A., Burlizzi, P., Carrasco, V., Chaikovsky, A., Comerón, A., D'Amico, G., Filioglou, M., Granados-Muñoz, M. J., Guerrero-Rascado, J. L., Ilic, L., Kokkalis, P., Maurizzi, A., Mona, L., Monti, F., Muñoz Porcar, C., Nicolae, D., Papayannis, A., Pappalardo, G., Pejanovic, G., Perreira, S., Perrone, M., Pietruczuk, A., Posyniak, M., Rocadenbosch, F., Rodríguez-Gómez, A., Sicard, M., Siomos, N., Szkop, A., Taradellas, E., Tsekeri, A., Vukovic, A., Wandinger, U., and Wagner, J.: A methodology for investigating dust model performance using synergistic EARLINET/AERONET dust concentration retrievals, Atmos. Meas. Tech., 8, 3577–3600, doi:10.5194/amt-8-3577-2015, 2015.

Bovchaliuk, V., Goloub, P., Podvin, T., Veselovskii, I., Tanré, D., Chaikovsky, A., Dubovik, O., Mortier, A., Lopatin, A., Korenskiy, M., and Victori, S.: Comparison of aerosol properties retrieved using GARRLiC, LIRIC, and Raman algorithms applied to multi-wavelength lidar and sun/sky-photometer data, Atmospheric Measurement Techniques, 9, 3391–3405, doi:10.5194/amt-9-3391-2016, 2016.

Bravo-Aranda, J. A., Titos, G., Granados-Muñoz, M., Guerrero-Rascado, J., Navas-Guzmán, F., Valenzuela, A., Lyamani, H., Olmo, F., Andrey, J., and Alados-Arboledas, L.: Study of mineral dust entrainment in the planetary boundary layer by lidar depolarisation technique, Tellus B: Chemical and Physical Meteorology, 67, 26 180, doi:10.3402/tellusb.v67.26180, 2015.

Burton, S. P., Ferrare, R. A., Hostetler, C. A., Hair, J. W., Rogers, R. R., Obland, M. D., Butler, C. F., Cook, A. L., Harper, D. B., and Froyd, K. D.: Aerosol classification using airborne High Spectral Resolution Lidar measurements–methodology and examples, Atmos. Meas. Tech., 5, 73–98, doi:10.5194/amt-5-73-2012, 2012.

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.

Burton, S. P., Vaughan, M. A., Ferrare, R. A., and Hostetler, C. A.: Separating mixtures of aerosol types in airborne High Spectral Resolution Lidar data, Atmos. Meas. Tech., 7, 419–436, doi:10.5194/amt-7-419-2014, 2014.

Burton, S. P., Hair, J. W., Kahnert, M., Ferrare, R. A., Hostetler, C. A., Cook, A. L., Harper, D. B., Berkoff, T. A., Seaman, S. T., Collins, J. E., Fenn, M. A., and Rogers, R. R.: Observations of the spectral dependence of linear particle depolarization ratio of aerosols using NASA Langley airborne High Spectral Resolution Lidar, Atmospheric Chemistry and Physics, 15, 13 453–13 473, doi:10.5194/acp-15-13453-2015, 2015.

Cattrall, C., Reagan, J., Thome, K., and Dubovik, O.: Variability of aerosol and spectral lidar and backscatter and extinction ratios of key aerosol types derived from selected Aerosol Robotic Network locations, J. Geophys. Res., 110, doi:10.1029/2004JD005124, 2005.

Chaikovsky, A., Dubovik, O., Holben, B., Bril, A., Goloub, P., Tanré, D., Pappalardo, G., Wandinger, U., Chaikovskaya, L., Denisov, S., Grudo, J., Lopatin, A., Karol, Y., Lapyonok, T., Amiridis, V., Ansmann, A., Apituley, A., Alados-Arboledas, L., Binietoglou, I., Boselli, A., D'Amico, G., Freudenthaler, V., Giles, D., Granados-Muñoz, M. J., Kokkalis, P., Nicolae, D., Oshchepkov, S., Papayannis, A., Perrone, M. R., Pietruczuk, A., Rocadenbosch, F., Sicard, M., Slutsker, I., Talianu, C., De Tomasi, F., Tsekeri, A., Wagner, J., and Wang, X.: Lidar-Radiometer Inversion Code (LIRIC) for the retrieval of vertical aerosol properties from combined lidar-radiometer data: development and distribution in EARLINET, Atmos. Meas. Tech., 9, 1181–1205, doi:10.5194/amt-9-1181-2016, 2016.

Córdoba-Jabonero, C., Sorribas, M., Guerrero-Rascado, J. L., Adame, J. A., Hernández, Y., Lyamani, H., Cachorro, V., Gil, M., Alados-Aroledas, L., Cuevas, E., and de la Morena, B.: Synergetic monitoring of Saharan dust plumes and potential impact on surface: a case study of dust transport from Canary Islands to Iberian Peninsula, Atmos. Chem. Phys., 11, 3067–3091, doi:10.5194/acp-11-3067-2011, 2011.

D'Amico, G., Amodeo, A., Baars, H., Binietoglou, I., Freudenthaler, V., Mattis, I., Wandinger, U., and Pappalardo, G.: EARLINET Single Calculus Chain–overview on methodology and strategy, Atmos. Meas. Tech., 8, 4891–4916, doi:10.5194/amt-8-4891-2015, 2015.

Dawson, K., Meskhidze, N., Josset, D., and Gassó, S.: Spaceborne observations of the lidar ratio of marine aerosols, Atmos. Chem. Phys., 15, 3241–3255, doi:10.5194/acp-15-3241-2015, 2015.

Freudenthaler, V., Esselborn, M., Wiegner, M., Heese, B., Tesche, M., Ansmann, A., Müller, D., Althausen, D., Wirth, M., Fix, A., Ehret, G., Knippertz, P., Toledano, C., Gasteiger, J., Garhammer, M., and Seefeldner, M.: Depolarization ratio profiling at several wavelengths in pure Saharan dust during SAMUM 2006, Tellus B, 61, 165–179, doi:10.1111/j.1600-0889.2008.00396.x, 2009.

Giannakaki, E., Balis, D., Amiridis, V., and Zerefos, C.: Optical properties of different aerosol types: seven years of combined Raman-elastic backscatter lidar measurements in Thessaloniki, Greece, Atmos. Meas. Tech., 3, 569–578, doi:10.5194/amt-3-569-2010, 2010.

Giannakaki, E., van Zyl, P. G., Müller, D., Balis, D., and Komppula, M.: Optical and microphysical characterization of aerosol layers over South Africa by means of multi-wavelength depolarization and Raman lidar measurements, Atmos. Chem. Phys., 16, 8109–8123, doi:10.5194/acp-16-8109-2016, 2016.

Ginoux, P., Chin, M., Tegen, I., Prospero, J. M., Holben, B., Dubovik, O., and Lin, S.-J.: Sources and distributions of dust aerosols simulated with the GOCART model, J. Geo. phys. Res., 20, 20 255–20 274, doi:10.1029/2000JD000053, 2001.

Ginoux, P., Prospero, J. M., Gill, T. E., Hsu, N. C., and Zhao, M.: Global-scale attribution of anthropogenic and natural dust sources and their emission rates based on MODIS Deep Blue aerosol products, Reviews of Geophysics, 50, doi:10.1029/2012RG000388, rG3005, 2012.

Granados-Muñoz, M. J., Bravo-Aranda, J. A., Baumgardner, D., Guerrero-Rascado, J. L., Pérez-Ramírez, D., Navas-Guzmán, F., Veselovskii, I., Lyamani, H., Valenzuela, A., Olmo, F. J., Titos, G., Andrey, J., Chaikovsky, A., Dubovik, O., Gil-Ojeda, M., and Alados-Arboledas, L.: A comparative study of aerosol microphysical properties retrieved from ground-based remote sensing and aircraft in situ measurements during a Saharan dust event, Atmospheric Measurement Techniques, 9, 1113–1133, doi:10.5194/amt-9-1113-2016, 2016a.

Granados-Muñoz, M. J., Navas-Guzmán, F., Guerrero-Rascado, J. L., Bravo-Aranda, J. A., Binietoglou, I., Pereira, S. N., Basart, S., Baldasano, J. M., Belegante, L., Chaikovsky, A., Comerón, A., D'Amico, G., Dubovik, O., Ilic, L., Kokkalis, P., Muñoz Porcar, C., Nickovic, S., Nicolae, D., Olmo, F. J., Papayannis, A., Pappalardo, G., Rodríguez, A., Schepanski, K., Sicard, M., Vukovic, A., Wandinger, U., Dulac, F., and Alados-Arboledas, L.: Profiling of aerosol microphysical properties at several EARLINET/AERONET sites during the July 2012 ChArMEx/EMEP campaign, Atmospheric Chemistry and Physics, 16, 7043–7066, doi:10.5194/acp-16-7043-2016, 2016b.

Groß, S., Tesche, M., Freudenthaler, V., Toledano, C., Wiegner, M., 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 lidar measurements during SAMUM 2, Tellus B, 63, 706–724, doi:10.3402/tellusb.v63i4.16369, 2011.

Groß, S., Esselborn, M., Weinzierl, B., Wirth, M., Fix, A., and Petzold, A.: Aerosol classification by airborne high spectral resolution lidar observations, Atmospheric Chemistry and Physics, 13, 2487–2505, doi:10.5194/acp-13-2487-2013, 2013.

Groß, S., Gasteiger, J., Freudenthaler, V., Müller, T., Sauer, D., Toledano, C., and Ansmann, A.: Saharan dust contribution to the Caribbean summertime boundary layer–a lidar study during SALTRACE, Atmos. Chem. Phys., 16, 11 535–11 546, doi:10.5194/acp-16-11535-2016, 2016.

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, Atmospheric Environment, 42, 2668–2681, doi:10.1016/j.atmosenv.2007.12.062, 2008.

Guerrero-Rascado, J. L., Olmo, F. J., Avilés-Rodríguez, I., Navas-Guzmán, F., Pérez-Ramírez, D., Lyamani, H., and Alados Arboledas, L.: Extreme Saharan dust event over the southern Iberian Peninsula in september 2007: active and passive remote sensing from surface and satellite, Atmospheric Chemistry and Physics, 9, 8453–8469, doi:10.5194/acp-9-8453-2009, 2009.

Hamill, P., Giordano, M., Ward, C., Giles, D., and Holben, B.: An AERONET-based aerosol classification using the Mahalanobis distance, Atmos. Env., 140, 213–233, doi:10.1016/j.atmosenv.2016.06.002, 2016.

Heese, B., Althausen, D., Baars, H., Bohlmann, S., and Deng, R.: Aerosol Properties over Southeastern China from Multi-Wavelength Raman and Depolarization Lidar Measurements, EPJ Web of Conferences, 119, 23 018, doi:10.1051/epjconf/201611923018, 2016.

Holben, B., Eck, T., Slutsker, I., Tanré, D., Buis, J., Setzer, A., Vermote, E., Reagan, J., Kaufman, Y., Nakajima, T., Lavenu, F., Jankowiak, I., and Smirnov, A.: AERONET—A Federated Instrument Network and Data Archive for Aerosol Characterization, Remote Sensing of Environment, 66, 0034–4257, doi:10.1016/S0034-4257(98)00031-5, 1998.

Iarlori, M., Madonna, F., Rizi, V., Trickl, T., and Amodeo, A.: Effective resolution concepts for lidar observations, Atmospheric Measurement Techniques, 8, 5157–5176, doi:10.5194/amt-8-5157-2015, 2015.

Kim, M.-H., Kim, S.-W., Yoon, S.-C., and Omar, A. H.: Comparison of aerosol optical depth between CALIOP and MODIS-Aqua for CALIOP aerosol subtypes over the ocean, Journal of Geophysical Research: Atmospheres, 118, 13,241–13,252, doi:10.1002/2013JD019527, 2013.

Kokkalis, P., Papayannis, A., Amiridis, V., Mamouri, R. E., Veselovskii, I., Kolgotin, A., Tsaknakis, G., Kristiansen, N. I., Stohl, A., and Mona, L.: Optical, microphysical, mass and geometrical properties of aged volcanic particles observed over Athens, Greece, during the Eyjafjallajökull eruption in April 2010 through synergy of Raman lidar and sunphotometer measurements, Atmospheric Chemistry and Physics, 13, 9303–9320, doi:10.5194/acp-13-9303-2013, 2013.

Liu, D., Wang, Z., Liu, Z., Winker, D., and Trepte, C.: A height resolved global view of dust aerosols from the first year CALIPSO lidar measurements, J. Geophys. Res., 113, doi:10.1029/2007JD009776, 2008.

Madonna, F., Amato, F., Vande Hey, J., and Pappalardo, G.: Ceilometer aerosol profiling versus Raman lidar in the frame of the INTERACT campaign of ACTRIS, Atmospheric Measurement Techniques, 8, 2207–2223, doi:10.5194/amt-8-2207-2015, 2015.

Mahalanobis, P. C.: On the generalized distance in statistics, in: Proceedings of the National Institute of Science of India, pp. 49–55, 1936.

Mahowald, N., Albani, S., Kok, J. F., Engelstaeder, S., Scanza, R., Ward, D. S., and Flanner, M. G.: The size distribution of desert dust aerosols and its impact on the Earth system, Aeolian Research, 15, 53–71, doi:10.1016/j.aeolia.2013.09.002, 2014.

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, Atmospheric Measurement Techniques, 2, 513–522, doi:10.5194/amt-2-513-2009, 2009.

Mamouri, R. E., Ansmann, A., Nisantzi, A., Kokkalis, P., Schwarz, A., and Hadjimitsis, D.: Low Arabian dust extinction-to-backscatter ratio, Geophysical Research Letters, 40, 4762–4766, doi:10.1002/grl.50898, 2013.

Mattis, I., Siefert, P., Müller, D., Tesche, M., Hiebsch, A., Kanitz, T., Schmidt, J., Finger, F., Wandinger, U., and Ansmann, A.: Volcanic aerosol layers observed with multiwavelength Raman lidar over central Europe in 2008-2009, Journal of Geophysical Research: Atmospheres, 115, n/a–n/a, doi:10.1029/2009JD013472, http://dx.doi.org/10.1029/2009JD013472, d00L04, 2010.

Mona, L. and Marenco, F.: Volcanic Ash: Hazard Observation, chap. Lidar Observations of Volcanic Particles, pp. 161–173, Elsevier, 2016.

Mona, L., Amodeo, A., Pandolfi, M., and Pappalardo, G.: Saharan dust intrusions in the Mediterranean area: Three years of Raman lidar measurements, Journal of Geophysical Research: Atmospheres, 111, doi:10.1029/2005JD006569, 2006.

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, Atmospheric Chemistry and Physics, 9, 7213–7228, doi:10.5194/acp-9-7213-2009, 2009.

Mona, L., Amodeo, A., D'Amico, G., Giunta, A., Madonna, F., and Pappalardo, G.: Multi-wavelength Raman lidar observations of the Eyjafjallajökull volcanic cloud over Potenza, southern Italy, Atmos. Chem. Phys., 12, 2229–2244, doi:10.5194/acp-12-2229-2012, 2012a.

Mona, L., Liu, Z., Müller, D., Omar, A., Papayannis, A., Sugimoto, N., Pappalardo, G., and Vaughan, M.: Measurements for Desert Dust Characterization: An Overview, Adv. Meteorol., 2012, pp. 36, doi:10.1155/2012/356265, 2012b.

Mona, L., Papagiannopoulos, N., Basart, S., Baldasano, J., Binietoglou, I., Cornacchia, C., and Pappalardo, G.: EARLINET dust observations vs. BSC-DREAM8b modeled profiles: 12-year-long systematic comparison at Potenza, Italy, Atmospheric Chemistry and Physics, 14, 8781–8793, doi:10.5194/acp-14-8781-2014, 2014.

Müller, D., Mattis, I., Ansmann, A., Wehner, B., Althausen, D., Wandinger, U., and Dubovik, O.: Closure study on optical and microphysical properties of a mixed urban and Arctic haze air mass observed with Raman lidar and Sun photometer, Journal of Geophysical Research: Atmospheres, 109, doi:10.1029/2003JD004200, 2004.

Müller, D., Mattis, I., Wandinger, U., Ansmann, A., Althausen, D., and Stohl, A.: Raman lidar observations of aged Siberian and Canadian forest fire smoke in the free troposphere over Germany in 2003: Microphysical particle characterization, J. Geophys. Res., 110, doi:10.1029/2004JD005756, 2005.

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

Müller, D., Mattis, I., Ansmann, A., Wandinger, U., Ritter, C., and Kaiser, D.: Multiwavelength Raman lidar observations of particle growth during long-range transport of forest-fire smoke in the free troposphere, Geophys. Res. Lett., 34, doi:10.1029/2006GL027936, 2007b.

Müller, D., Heinold, B., Tesche, M., Tegen, I., Althausen, D., Alados-Arboledas, L., Amiridis, V., Amodeo, A., Ansmann, A., Balis, D., Comeron, A., D'Amico, G., Gerasopoulos, E., Guerrero-Rascado, J. L., Freudenthaler, V., Giannakaki, E., Heese, B., Iarlori, M., Knippertz, P., Mamouri, R. E., Mona, L., Papayannis, A., Pappalardo, G., Perrone, R.-M., Pisani, G., Rizi, V., Sicard, M., Spinelli, N., Tafuro, A., and Wiegner, M.: EARLINET observations of the 14–22-May long-range dust transport event during SAMUM 2006: validation of results from dust transport modelling, Tellus B, 61, 325–339, doi:10.1111/j.1600-0889.2008.00400.x, 2009.

Navas-Guzmán, F., Müller, D., Bravo-Aranda, J. A., Guerrero-Rascado, J. L., Granados-Muñoz, M. J., Pérez-Ramírez, D., Olmo, F. J., and Alados-Arboledas, L.: Eruption of the Eyjafjallajökull Volcano in spring 2010: Multiwavelength Raman lidar measurements of sulphate particles in the lower troposphere, Journal of Geophysical Research: Atmospheres, 118, 1804–1813, doi:10.1002/jgrd.50116, http://dx.doi.org/10.1002/jgrd.50116, 2013.

Navas-Guzmán, N., Bravo-Aranda, J. A., Guerrero-Rascado, J. L., Granados-Muñoz, M. J., and Alados-Arboledas, L.: Statistical analysis of aerosol optical properties retrieved by Raman lidar over Southeastern Spain, Tellus B: Chemical and Physical Meteorology, 65, 21 234, doi:10.3402/tellusb.v65i0.21234, 2013.

Nepomuceno Pereira, S., Preißler, J., Guerrero-Rascado, J. L., Silva, A. M., and Wagner, F.: Forest Fire Smoke Layers Observed in the Free Troposphere over Portugal with a Multiwavelength Raman Lidar: Optical and Microphysical Properties, Scientific World Journal, 2014, doi:10.1155/2014/42183, 2014.

Nicolae, D., Nemuc, A., Müller, D., Talianu, C., Vasilescu, J., Belegante, L., and Kolgotin, A.: Characterization of fresh and aged biomass burning events using multiwavelength Raman lidar and mass spectrometry, J. Geophys. Res. Atmos., 118, 2956–2965, doi:10.1002/jgrd.50324, 2013.

Nicolae, D., Vasilescu, J., Talianu, C., and Dandocsi, A.: Independent Retrieval of Aerosol Type From Lidar, EPJ Web of Conferences, 119, 18 002, doi:10.1051/epjconf/201611918002, 2016.

Nisantzi, A., Mamouri, R. E., Ansmann, A., Schuster, G. L., and Hadjimitsis, D. G.: Middle East versus Saharan dust extinction-to-backscatter ratios, Atmospheric Chemistry and Physics, 15, 7071–7084, doi:10.5194/acp-15-7071-2015, 2015.

O'Dowd, C. and de Leeuw, G.: Marine aerosol production: a review of the current knowledge, Phil. Tran. of the Roy. Soc. A, doi:10.1098/rsta.2007.2043, 2007.

Omar, A., Winker, D., Kittaka, C., Vaughan, M., Liu, Z., Hu, 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.

Ortiz-Amezcua, P., Guerrero-Rascado, J. L., Granados-Muñoz, M. J., Benavent-Oltra, J. A., Böckmann, C., Samaras, S., Stachlewska, I. S., Janicka, Ł., Baars, H., Bohlmann, S., and Alados-Arboledas, L.: Microphysical characterization of long-range transported biomass burning particles from North America at three EARLINET stations, Atmospheric Chemistry and Physics, 17, 5931–5946, doi:10.5194/acp-17-5931-2017, 2017.

Papagiannopoulos, N., Mona, L., Alados-Arboledas, L., Amiridis, V., Baars, H., Binietoglou, I., Bortoli, D., D'Amico, G., Giunta, A., Guerrero-Rascado, J. L., Schwarz, A., Perreira, S., Spinelli, N., Wandinger, U., Wang, X., and Pappalardo, G.: CALIPSO climatological products: evaluation and suggestions from EARLINET, Atmos. Chem. Phys., 16, 2341–2357, doi:10.5194/acp-16-2341-2016, 2016a.

Papagiannopoulos, N., Mona, L., and Pappalardo, G.: Aerosol classification using EARLINET measurements for an intensive observational period, vol. 18, European Geosciences Union, General Assembly, 2016b.

Papayannis, A., Amiridis, V., Mona, L., Tsaknakis, G., Balis, D., Bösenberg, J., Chaikovski, A., De Tomasi, F., Grigorov, I., Mattis, I., Mitev, V., Müller, D., Nickovic, S., Pérez, C., Pietruczuk, A., Pisani, G., Ravetta, F., Rizi, V., Sicard, M., Trickl, T., Wiegner, M., Gerding, M., Mamouri, R. E., D'Amico, G., and Pappalardo, G.: Systematic lidar observations of Saharan dust over Europe in the frame of EARLINET (2000-2002), J. Geophys. Res., 113, doi:10.1029/2007JD009028, 2008.

Papayannis, A., Mamouri, R. E., Amiridis, V., Giannakaki, E., Veselovskii, I., Kokkalis, P., Tsaknakis, G., Balis, D., Kristiansen, N., Stohl, A., Korenskiy, M., Allakhverdiev, K., Huseyinoglu, M., and Baykara, T.: Optical properties and vertical extension of aged ash layers over the Eastern Mediterranean as observed by Raman lidars during the Eyjafjallajökull eruption in May 2010, Atmospheric Environment, 48, 56–65, doi:10.1016/j.atmosenv.2011.08.037, 2012.

Papayannis, A., Nicolae, D., Kokkalis, P., Binietoglou, I., Talianu, C., Belegante, L., Tsaknakis, G., Cazacu, M., Vetres, I., and Ilic, L.: Optical, size and mass properties of mixed type aerosols in Greece and Romania as observed by synergy of lidar and sunphotometers in combination with model simulations: A case study, Science of The Total Environment, 500-501, 277–297, doi:10.1016/j.scitotenv.2014.08.101, 2014.

Papayannis, A., Argyrouli, A., Müller, D., Tsaknakis, G., Kokkalis, P., Binietoglou, I., Kazadzis, S., Solomos, S., and Amiridis, V.: Tropospheric Vertical Profiles of Aerosol Optical, Microphysical and Concentration Properties in the Frame of the Hygra-CD Campaign (Athens, Greece 2014): A Case Study of Long-Range Transport of Mixed Aerosols, EPJ Web of Conferences, 119, 23 016, doi:10.1051/epjconf/201611923016, 2016.

Pappalardo, G., Amodeo, A., Mona, L., Pandolfi, M., Pergola, N., and Cuomo, V.: Raman lidar observations of aerosol emitted during the 2002 Etna eruption, Geophysical Research Letters, 31, doi:10.1029/2003GL019073, 2004a.

Pappalardo, G., Amodeo, A., Pandolfi, M., Wandinger, U., Ansmann, A., Bösenberg, J., Matthias, V., Amiridis, V., De Tomasi, F., Frioud, M., Iarlori, M., Komguem, L., Papayannis, A., Rocadenbosch, F., and Wang, X.: Aerosol lidar intercomparison in the framework of the EARLINET project. 3. Ramanlidar algorithm for aerosol extinction, backscatter, and lidar ratio, Appl. Opt., 43, 5370–5385, doi:10.1364/AO.43.005370, 2004b.

Pappalardo, G., Mona, L., D'Amico, G., Wandinger, U., Adam, M., Amodeo, A., Ansmann, A., Apituley, A., Alados Arboledas, L., Balis, D., Boselli, A., Bravo-Aranda, J. A., Chaikovsky, A., Comerón, A., Cuesta, J., De Tomasi, F., Freudenthaler, V., Gausa, M., Giannakaki, E., Giehl, H., Giunta, A., Grigorov, I., Groß, S., Haeffelin, M., Hiebsch, A., Iarlori, M., Lange, D., Linné, H., Madonna, F., Mattis, I., Mamouri, R. E., McAuliffe, M. A. P., Mitev, V., Molero, F., Navas-Guzmán, F., Nicolae, D., Papayannis, A., Perrone, M. R., Pietras, C., Pietruczuk, A., Pisani, G., Preißler, J., Pujadas, M., Rizi, V., Ruth, A. A., Schmidt, J., Schnell, F., Seifert, P., Serikov, I., Sicard, M., Simeonov, V., Spinelli, N., Stebel, K., Tesche, M., Trickl, T., Wang, X., Wagner, F., Wiegner, M., and Wilson, K. M.: Four-dimensional distribution of the 2010 Eyjafjallajökull volcanic cloud over Europe observed by EARLINET, Atmos. Chem. Phys., 13, 4429–4450, doi:10.5194/acp-13-4429-2013, 2013.

Pappalardo, G., Amodeo, A., Apituley, A., Comerón, A., Freudenthaler, V., Linné, H., Ansmann, A., Bösenberg, J., D'Amico, G., Mattis, I., Mona, L., Wandinger, U., Amiridis, V., Alados-Arboledas, L., Nicolae, D., and Wiegner, M.: EARLINET: towards an advanced sustainable European aerosol lidar network, Atmos. Meas. Tech., 7, 2389–2409, doi:10.5194/amt-7-2389-2014, 2014.

Prata, F., Woodhouse, M., Huppert, H. E., Prata, A., Thordarson, T., and Carn, S.: Atmospheric processes affecting the separation of volcanic ash and SO_2 in volcanic eruptions: inferences from the May 2011 Grímsvötn eruption, Atmospheric Chemistry and Physics, 17, 10 709–10 732, doi:10.5194/acp-17-10709-2017, 2017.

Preißler, J., Wagner, F., Pereira, S. N., and Guerrero-Rascado, J. L.: Multi-instrumental observation of an exceptionally strong Saharan dust outbreak over Portugal, Journal of Geophysical Research: Atmospheres, 116, doi:10.1029/2011JD016527, 2011.

Preißler, J., Wagner, F., Guerrero-Rascado, J. L., and Silva, A. M.: Two years of free-tropospheric aerosol layers observed over Portugal by lidar, J. Geophys. Res. Atmos., 118, 3676–3686, doi:10.1002/jgrd.50350, 2013.

Rencher, A.: Methods of Multivariate Analysis, Wiley-Interscience, 2002.

Rogers, R. R., Vaughan, M. A., Hostetler, C. A., Burton, S. P., Ferrare, R. A., Young, S. A., Hair, J. W., Obland, M. D., Harper, D. B., Cook, A. L., and Winker, D. M.: Looking through the haze: evaluating the CALIPSO level 2 aerosol optical depth using airborne high spectral resolution lidar data, Atmospheric Measurement Techniques, 7, 4317–4340, doi:10.5194/amt-7-4317-2014, 2014.

Russell, P. B., Kacenelenbogen, M., Livingston, J. M., Hasekamp, O. P., Burton., S. P., Schuster, G. L., Johnson, M. S., Knobelspiesse, K. D., Redemann, J., Ramachandran, S., and Holben, B.: A multiparameter aerosol classification method and its application to retrievals from spaceborne polarimetry, J. Geophys. Res. Atmos., 119, 9838–9863, doi:10.1002/2013JD021411, 2014.

Samaras, S., Nicolae, D., Böckmann, C., Vasilescu, J., Binietoglou, I., Labzovskii, L., Toanca, F., and Papayannis, A.: Using Raman-lidar-based Regularized Microphysical Retrievals and Aerosol Mass Spectrometer Measurements for the Characterization of Biomass Burning Aerosols, J. Comput. Phys., 299, 156–174, doi:10.1016/j.jcp.2015.06.045, 2015.

Sawamura, P., Vernier, J. P., Barnes, J. E., Berkoff, T. E., Welton, E. J., Alados-Arboledas, L., Navas-Guzmán, F., Pappalardo, G., Mona, L., Madonna, F., Lange, D., Sicard, M., Godin-Beekmann, S., Payen, G., Wang, Z., Hu, S., Tripathi, S. N., Cordoba-Jabonero, C., and Hoff, R. M.: Stratospheric AOD after the 2011 eruption of Nabro volcano measured by lidars over the Northern Hemisphere, Environmental Research Letters, 7, 034 013, doi:doi:10.1088/1748-9326/7/3/034013, http://stacks.iop.org/1748-9326/7/i=3/a=034013, 2012.

Schuster, G. L., Vaughan, M., MacDonnell, D., Su, W., Winker, D., Dubovik, O., Lapyonok, T., and Trepte, C.: Comparison of CALIPSO aerosol optical depth retrievals to AERONET measurements, and a climatology for the lidar ratio of dust, Atmos. Chem. Phys., 12, 7431–7452, doi:10.5194/acp-12-7431-2012, 2012.

Schwarz, A.: Aerosol typing over Europe and its benefits for the CALIPSO and EarthCARE missions - statistical analysis based on multiwavelength aerosol lidar measurements from ground-based EARLINET stations and comparison to spaceborne CALIPSO data, Ph.D. thesis, University of Leipzig, 2016.

Sicard, M., Guerrero-Rascado, J. L., Navas-Guzmán, F., Preißler, J., Molero, F., Tomás, S., Bravo-Aranda, J. A., Comerón, A., Rocadenbosch, F., Wagner, F., Pujadas, M., and Alados-Arboledas, L.: Monitoring of the Eyjafjallajökull volcanic aerosol plume over the Iberian Peninsula by means of four EARLINET lidar stations, Atmospheric Chemistry and Physics, 12, 3115–3130, doi:10.5194/acp-12-3115-2012, https://www.atmos-chem-phys.net/12/3115/2012/, 2012a.

Sicard, M., Mallet, M. García-Vizcaíno, D., Comerón, A., Rocadenbosch, F., Dubuisson, P., and Muñoz Porcar, C.: Intense dust and extremely fresh biomass burning outbreak in Barcelona, Spain: characterization of their optical properties and estimation of their direct radiative forcing, Environ. Res. Lett., 7, 6pp, doi:10.1088/1748-9326/7/3/034016, 2012b.

Sicard, M., D'Amico, G., Comerón, A., Mona, L., Alados-Arboledas, L., Amodeo, A., Baars, H., Baldasano, J. M., Belegante, L., Binietoglou, I., Bravo-Aranda, J. A., Fernández, A. J., Fréville, P., García-Vizcaíno, D., Giunta, A., Granados-Muñoz, M. J., Guerrero-Rascado, J. L., Hadjimitsis, D., Haefele, A., Hervo, M., Iarlori, M., Kokkalis, P., Lange, D., Mamouri, R. E., Mattis, I., Molero, F., Montoux, N., Muñoz, A., Muñoz Porcar, C., Navas-Guzmán, F., Nicolae, D., Nisantzi, A., Papagiannopoulos, N., Papayannis, A., Pereira, S., Preißler, J., Pujadas, M., Rizi, V., Rocadenbosch, F., Sellegri, K., Simeonov, V., Tsaknakis, G., Wagner, F., and Pappalardo, G.: EARLINET: potential operationality of a research network, Atmospheric Measurement Techniques, 8, 4587–4613, doi:10.5194/amt-8-4587-2015, 2015.

Stohl, A., Forster, C., Frank, A., Seibert, P., and Wotawa, G.: Technical note: The Lagrangian particle dispersion model FLEXPART version 6.2, Atmospheric Chemistry and Physics, 5, 2461–2474, doi:10.5194/acp-5-2461-2005, 2005.

Tanré, D., Bréon, F. M., Deuzé, J. L., Dubovik, O., Ducos, F., François, P., Goloub, P., Herman, M., Lifermann, A., and Waquet, F.: Remote sensing of aerosols by using polarized, directional and spectral measurements within the A-Train: the PARASOL mission, Atmospheric Measurement Techniques, 4, 1383–1395, doi:10.5194/amt-4-1383-2011, 2011.

Tesche, M., Müller, D., Groß, S., Ansmann, A., Althausen, D., Freudenthaler, V., Weinzierl, B., Veira, A., and Petzold, A.: Optical and microphysical properties of smoke over Cape Verde inferred from multiwavelength lidar measurements, Tellus B, 63, 677–694, doi:10.1111/j.1600-0889.2011.00549.x, 2011.

Tesche, M., Wandinger, U., Ansmann, A., Althausen, D., Müller, D., and Omar, A. H.: 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.

Valenzuela, A., Olmo, F. J., Lyamani, H., Antón, M., Quirantes, A., and Alados-Arboledas, L.: Classification of aerosol radiative properties during African desert dust intrusions over southeastern Spain by sector origins and cluster analysis, Journal of Geophysical Research: Atmospheres, 117, doi:10.1029/2011JD016885, d06214, 2012.

Valenzuela, A., Olmo, F. J., Lyamani, H., Granados-Muñoz, M. J., Antón, M., Guerrero-Rascado, J. L., Quirantes, A., Toledano, C., Perez-Ramírez, D., and Alados-Arboledas, L.: Aerosol transport over the western Mediterranean basin: Evidence of the contribution of fine particles to desert dust plumes over Alborán Island, Journal of Geophysical Research: Atmospheres, 119, 14,028–14,044, doi:10.1002/2014JD022044, 2014JD022044, 2014.

Veselovskii, I., Dubovik, O., Kolgotin, A., Lapyonok, T., Di Girolamo, P., Summa, D., Whiteman, D. N., Mishchenko, M., and Tanré, D.: Application of randomly oriented spheroids for retrieval of dust particle parameters from multiwavelength lidar measurements, J. Geophys. Res., 115, doi:10.1029/2010JD014139, 2010.

Wandinger, U., Baars, H., Engelmann, R., Hünerbein, A., Horn, S., Kanitz, T., Donovan, D., van Zadelhoff, G.-J., Daou, D., Fischer, J., von Bismarck, J., Filipitsch, F., Docter, N., Eisinger, M., Lajas, D., and Wehr, T.: HETEAC: The Aerosol Classification Model for EarthCARE, EPJ Web of Conferences, 119, 01 004, doi:10.1051/epjconf/201611901004, 2016.

Wang, X., Boselli, A., D'Avino, L., Pisani, G., Spinelli, N., Amodeo, A., Chaikovsky, A., Wiegner, M., Nickovic, S., Papayannis, A., Perrone, M., Rizi, V., Sauvage, L., and Stohl, A.: Volcanic dust characterization by EARLINET during Etna's eruptions in 2001-2002, Atmospheric Environment, 42, 893–905, doi:10.1016/j.atmosenv.2007.10.020, 2008.

Wiegner, M., Groß, S., Freudenthaler, V., Schnell, F., and Gasteiger, J.: The May/June 2008 Saharan dust event over Munich: Intensive aerosol parameters from lidar measurements, J. Geophys. Res., 116, doi:10.1029/2011JD016619, 2011.

Wiegner, M., Gasteiger, J., Groß, S., Schnell, F., Freudenthaler, V., and Forkel, R.: Characterization of the Eyjafjallajökull ash-plume: Potential of lidar remote sensing, Physics and Chemistry of the Earth, Parts A/B/C, 45–46, 79–86, doi:10.1016/j.pce.2011.01.006, 2012.

Wiegner, M., Madonna, F., Binietoglou, I., Forkel, R., Gasteiger, J., Geiß, A., Pappalardo, G., Schäfer, K., and Thomas, W.: What is the benefit of ceilometers for aerosol remote sensing? An answer from EARLINET, Atmospheric Measurement Techniques, 7, 1979–1997, doi:10.5194/amt-7-1979-2014, 2014.

Wilks, S.: Multivariate statistical outliers, Sankhya: The Indian Journal of Statistics, 4, 407–426, 1963.

Table 1. Number of classified aerosol layers adapted from Schwarz (2016). The mixtures category is comprised of two or more ~~(values in parentheses)~~ pure aerosol types.

Aerosol type	All analyzed	Only from $3\beta + 2\alpha$
Clean Continental (CC)	45	5
Polluted Continental (PC)	95	19
Dust (D)	41	6
Mixed Dust (MD)	56	9
Polluted Dust (PD)	14	3
Smoke (S)	24	7
Volcanic (V)	21	4
Mixtures	348 ~~(170)~~	35 ~~(11)~~
Total	644	88

Table 2. Reference dataset: mean type-dependent intensive properties along with the standard deviation.

Type	$\kappa_\beta(355, 1064)$	$\kappa_\beta(532, 1064)$	$\kappa_\beta(355, 532)$	$\kappa_\alpha(355, 532)$	S_{aer}^{355} [sr]	S_{aer}^{532} [sr]	# ~~Samples~~ Layers
CC	1.0 ± 0.2	1.0 ± 0.3	1.3 ± 0.3	1.7 ± 0.6	50 ± 8	41 ± 6	9
PC	1.3 ± 0.3	1.3 ± 0.2	1.4 ± 0.6	1.7 ± 0.5	69 ± 12	63 ± 13	16
D	0.4 ± 0.1	0.4 ± 0.1	0.3 ± 0.2	0.3 ± 0.4	58 ± 12	55 ± 7	9
MD	0.5 ± 0.2	0.4 ± 0.3	0.7 ± 0.3	0.5 ± 0.3	42 ± 4	47 ± 6	10
PD	0.9 ± 0.3	0.8 ± 0.1	1.0 ± 0.5	0.6 ± 0.2	54 ± 8	64 ± 9	5
MM	0.8 ± 0.1	0.8 ± 0.2	1.0 ± 0.3	0.9 ± 0.3	25 ± 7	24 ± 8	8
S	1.3 ± 0.1	1.3 ± 0.1	1.2 ± 0.3	1.3 ± 0.3	81 ± 16	78 ± 11	7
V	0.1 ± 0.1	0.4 ± 0.3	0.2 ± 0.3	0.2 ± 0.3	50 ± 11	48 ± 13	5

Table 3. Aerosol types that constitute the classes investigated. CC stands for Clean Continental, PC stands for polluted continental, D stands for dust, MD stands for mixed dust, PD stands for polluted dust, MM stands for mixed marine, S stands for smoke, and V stands for volcanic particles.

# Types	Groups of aerosol types							
8	D	V	MD	PD	CC	MM	PC	S
7^a	D+V	MD	PD	CC	MM	PC	S	
7^b	D	V	MD	PD	CC	MM	PC+S	
6	D+V	MD	PD	CC	MM	PC+S		
5	D+V+MD+PD	CC	MM	PC	S			
4	D+V+MD+PD	CC	MM	PC+S				

Table 4. The mean and standard deviation of the particle depolarization ratio used for the pre-specified classes and the corresponding bibliographic references.

Type	δ_{aer}^{532}	References
Clean Continental	0.04 ± 0.02	Burton et al. (2013)
Polluted Continental	0.05 ± 0.03	Burton et al. (2013)
Dust	0.30 ± 0.01	Groß et al. (2011)
Mixed Dust	0.15 ± 0.02	Groß et al. (2016)
Polluted Dust	0.20 ± 0.05	Burton et al. (2013)
Marine	0.03 ± 0.01	Groß et al. (2013)
Smoke	0.10 ± 0.04	Burton et al. (2013)
Volcanic	0.33 ± 0.03	Pappalardo et al. (2013)

Table 5. Testing dataset: mean type-dependent intensive properties along with the standard deviation.

Type	$\kappa_\beta(355, 1064)$	$\kappa_\beta(532, 1064)$	$\kappa_\beta(355, 532)$	$\kappa_\alpha(355, 532)$	S_{aer}^{355} [sr]	S_{aer}^{532} [sr]	# ~~Samples~~ Layers
CC	1.2 ± 0.4	1.6 ± 0.4	1.3 ± 0.2	1.6 ± 0.4	43 ± 5	38 ± 6	12
PC	1.3 ± 0.4	1.4 ± 0.3	1.3 ± 0.3	1.2 ± 0.3	52 ± 6	56 ± 8	8
D	0.3 ± 0.3	0.2 ± 0.3	0.3 ± 0.2	0.0 ± 0.2	54 ± 11	54 ± 9	13
MM	0.8 ± 0.2	1.2 ± 0.5	0.9 ± 0.3	0.9 ± 0.3	27 ± 9	24 ± 8	8
S	1.6 ± 0.4	1.6 ± 0.5	1.5 ± 0.3	1.6 ± 0.4	54 ± 9	61 ± 6	6

Table 6. Recall (R), and precision (P) estimated from the classification matrices for 8, 7^a, 7^b, 6, 5, and 4 classes. The values between parentheses correspond to the number of layers passed the screening criteria.

Types	R [%]	P [%]	Types	R [%]	P [%]	Types	R [%]	P [%]
8 classes (29/47)			**7^a classes (29/47)**			**7^b classes (34/47)**		
CC	100	82	CC	100	89	CC	100	90
D	30	100	D+V	55	100	D	27	100
MM	100	100	MM	100	100	MM	100	100
PC	25	33	PC	50	50	PC+~~C~~S	78	100
S	0	-	S	0	-			
6 classes (33/47)			**5 classes (35/47)**			**4 classes (39/47)**		
CC	100	80	CC	100	82	CC	100	85
D+V	58	100	D+V+PD+MD	100	87	D+V+PD+MD	100	87
MM	100	100	MM	100	100	MM	100	100
PC+S	63	100	PC	20	33	PC+S	56	100
			~~CC~~S	0	-			

Table 7. Recall (R) and precision (P) estimated from the classification matrices for 8, 7^a, 7^b, 6, 5, and 4 classes when particle linear depolarization ratio measurements are available. The values between parentheses correspond to the number of layers passed the screening criteria.

Types	R [%]	P [%]	Types	R [%]	P [%]	Types	R [%]	P [%]
8 classes (14/21)			**7^a classes (13/21)**			**7^b classes (16/21)**		
CC	75	100	CC	100	100	CC	~~100~~ 75	100
D	~~67~~ 88	100	D+V	~~80~~ 88	100	D	~~67~~ 88	100
MM	-	-	MM	-	-	MM	-	-
PC	50	~~100~~ 50	PC	100	~~50~~ 75	PC+S	~~100~~ 75	~~100~~ 75
S	~~0~~ -	-	S	0	-			
6 classes (16/21)			**5 classes (13/21)**			**4 classes (15/21)**		
CC	~~100~~ 75	100	CC	100	100	CC	~~100~~ 75	100
D+V	88	100	D+V+PD+MD	100	~~89~~ 100	D+V+PD+MD	100	~~89~~ 80
MM	-	-	MM	-	-	MM	-	-
PC+S	~~100~~ 75	~~100~~ 75	PC	~~50~~ 0	~~50~~ -	PC+S	~~67~~ 33	~~100~~ 50
			S	~~0~~ -	-			

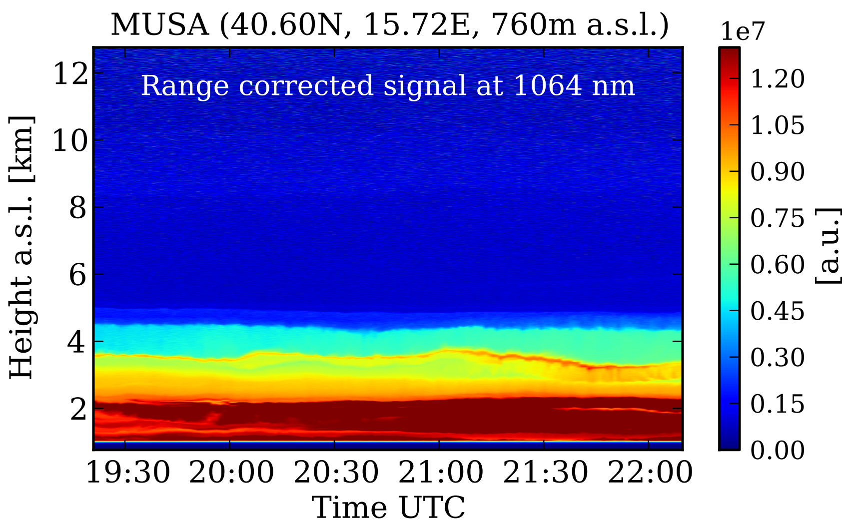

Figure 1. Temporal evolution of the 1064- nm range corrected lidar signal obtained with the MUSA system in Potenza on 14/07/2011, 19:20–22:10 UTC.

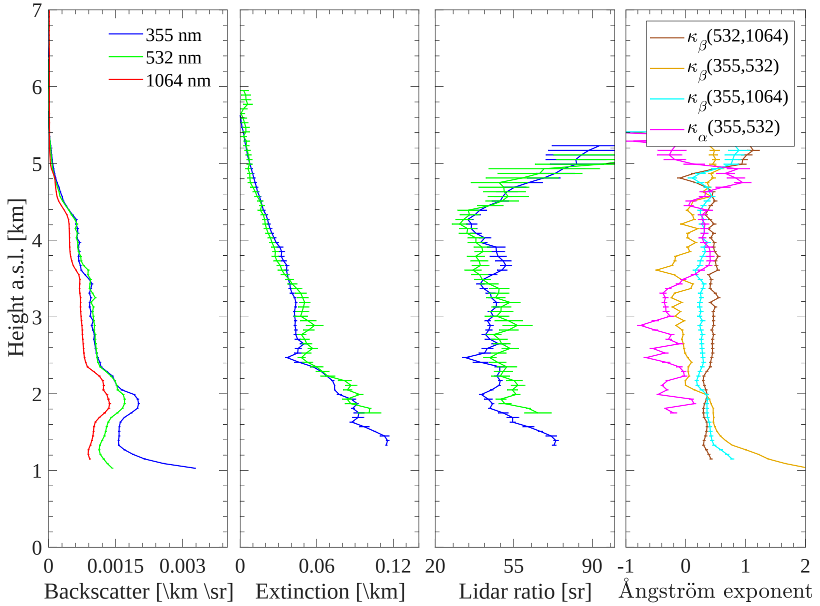

Figure 2. Optical profiles measured in Potenza, on 14 July 2011, 19:20–22:10 UTC with a multiwavelength Raman lidar. The error bars correspond to the standard deviation.

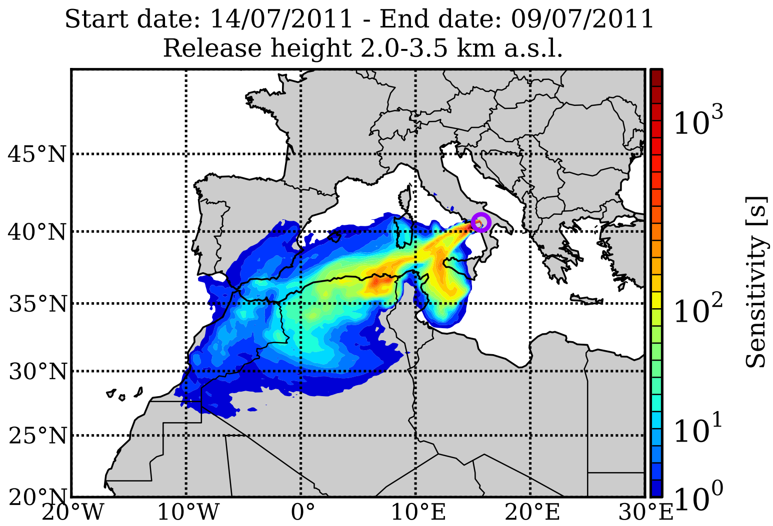

Figure 3. FLEXPART footprint for the airmass travelling below 2 km height and arriving at Potenza between ~~1.5–5~~ 2.0–3.5 km at 22:00 UTC on 14 July 2011. The colors are coded with respect to the logarithm of the integrated residence time in a grid box in seconds for a 5-day integration time.

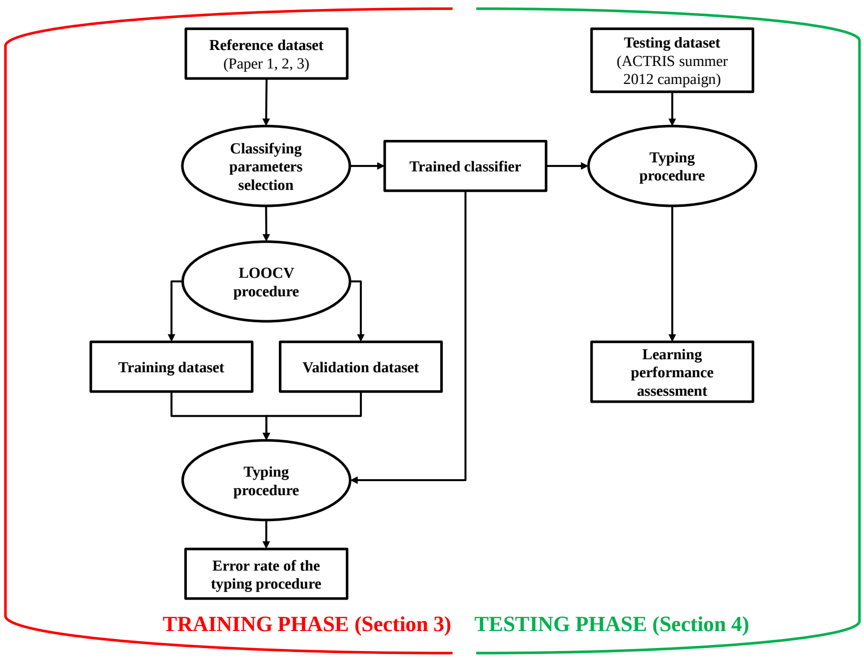

Figure 4. Flowchart of the methodology. First, well characterized aerosol layers are grouped into meaningful classes that represent the reference dataset: Paper 1 (Papagiannopoulos et al., 2016a), Paper 2 (Pappalardo et al., 2013), Paper 3 (Schwarz, 2016). Second, an analysis is performed to determine the best performing classifying parameters among the available intensive parameters. Third, based on the reference dataset the selected classifier is validated using the Leave-one-out cross validation (LOOCV) procedure in order to ensure correct aerosol type separation. Finally, the trained typing algorithm is applied to an independent and manually typed dataset (the testing dataset) for the assessment of the algorithm performance. Note that both phases have been applied with and without the depolarization ratio.

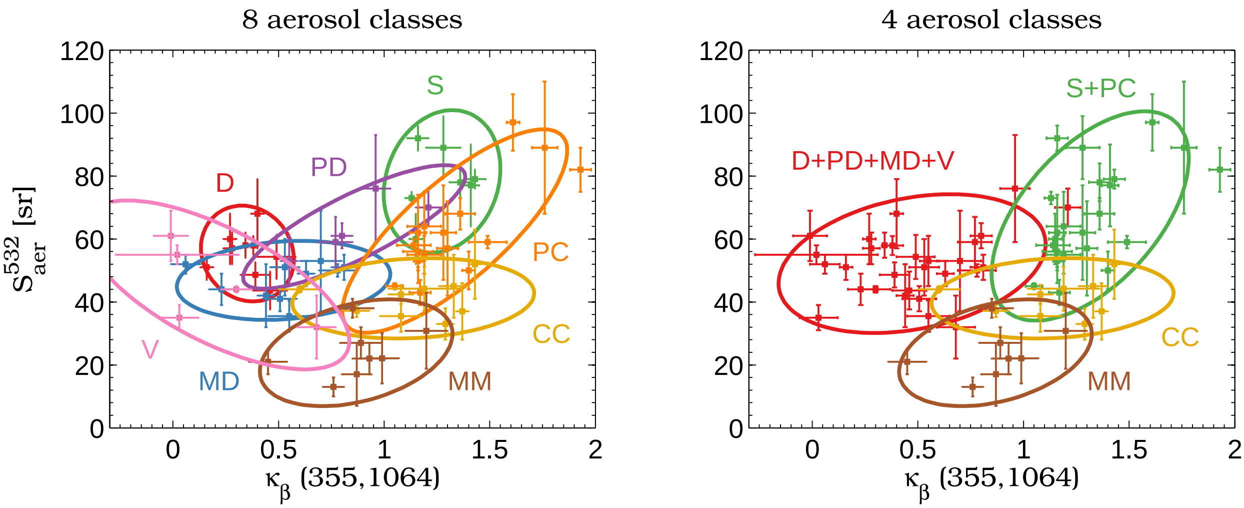

Figure 5. Colored pre-specified classes and 90 % confidence ellipses for 8 and 4 aerosol classes. The error bars correspond to the standard deviation of the selected mean intensive properties. CC stands for Clean Continental, D stands for dust, MD stands for mixed dust, MM stands for mixed marine, PD stands for polluted dust, PC stands for polluted continental, S stands for smoke, and V stands for volcanic particles.

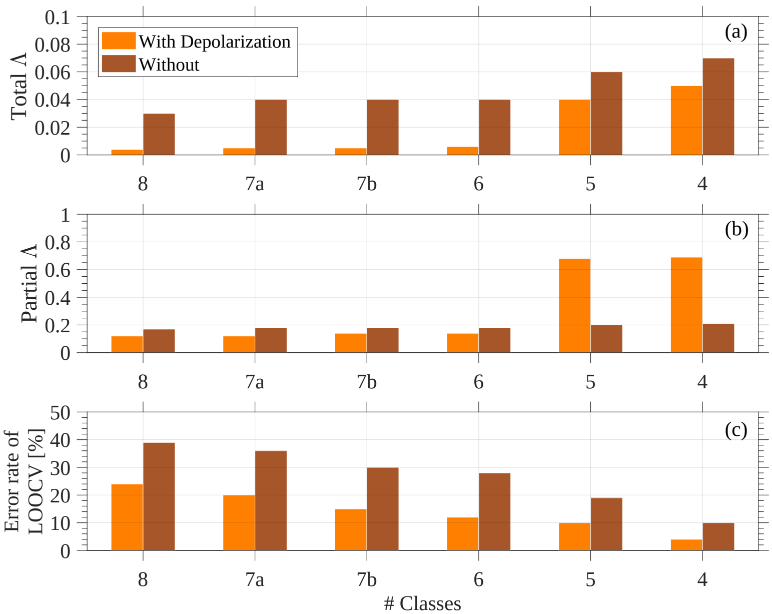

Figure 6. Bar plots showing a) the total Λ, b) the partial Λ, and c)error rate of LOOCV when comparing the training of the algorithm with (i.e., $S_{aer}^{532}/S_{aer}^{355}$, S_{aer}^{532}, and $\cancel{\kappa_\beta}\kappa_\beta(355, 1064)$) and without (i.e., δ_{aer}^{532}, S_{aer}^{532}, and $\cancel{\kappa_\beta}\kappa_\beta(355, 1064)$) particle linear depolarization values. For the partial Λ, the brown bars correspond to the backscatter-related Ångström exponent and orange one to the particle linear depolarization ratio because because they represent the most significant classifying parameter of the classification.

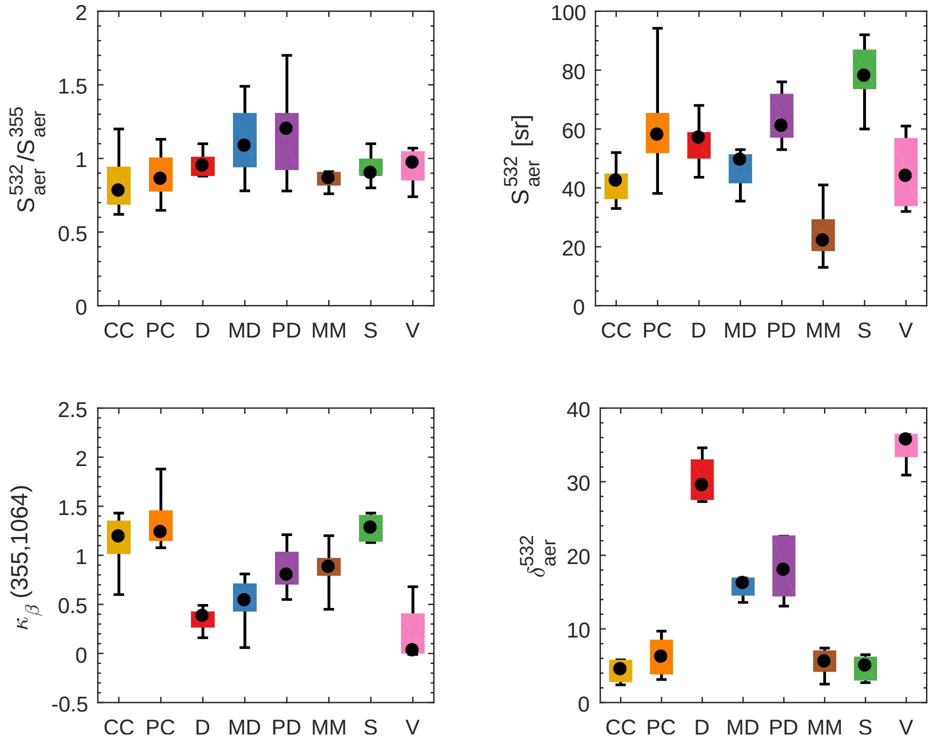

Figure 7. Bar plots show the median (horizontal line), 25-75 percentile (box) and 5-95 percentile (whisker) of the four classifying parameters: δ^{532}_{aer}, $\kappa_\beta(355, 1064)$, S^{532}_{aer}, and $S^{532}_{aer}/S^{355}_{aer}$. CC stands for Clean Continental, D stands for dust, MD stands for mixed dust, MM stands for mixed marine, PD stands for polluted dust, PC stands for polluted continental, S stands for smoke, and V stands for volcanic particles.

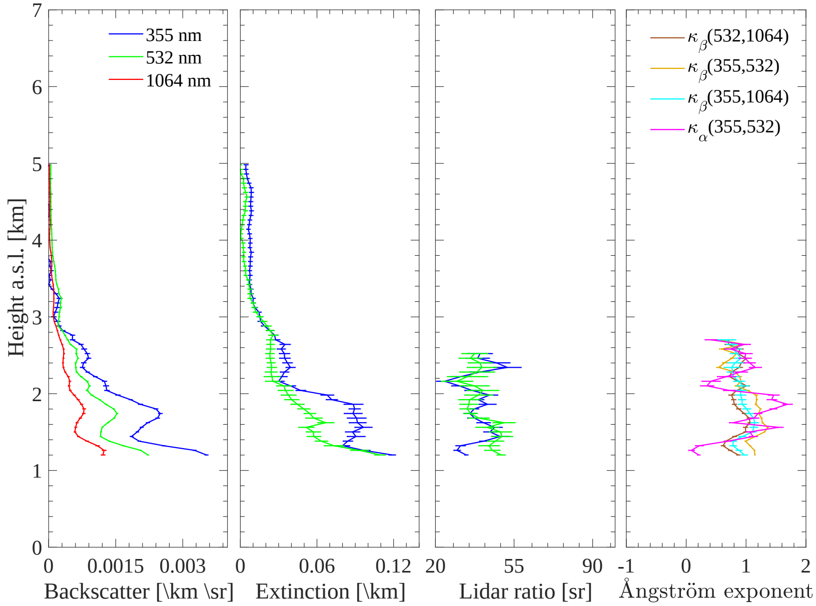

Figure 8. Optical profiles measured at Athens, on 22 May 2014, 20:28–21:28 UTC with a multiwavelength Raman lidar. The error bars correspond to the standard deviation. The effective resolution of the extinction coefficient profiles varied from 240 m to 780 m using the method described in Pappalardo et al. (2004b).

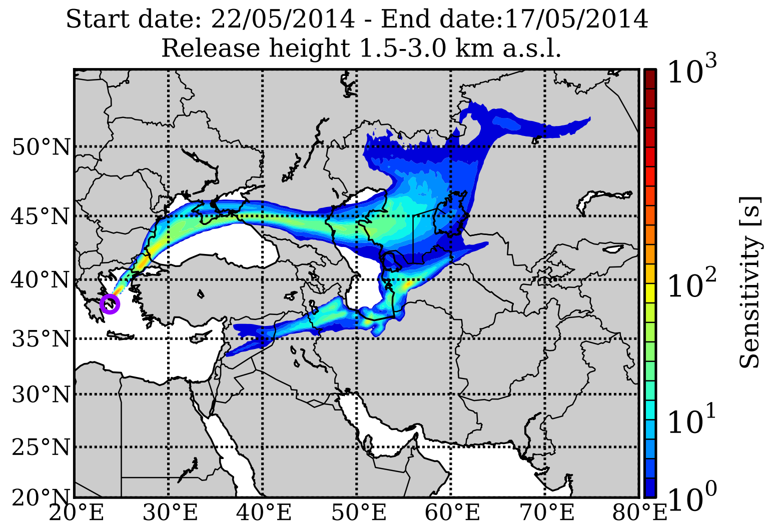

Figure 9. FLEXPART footprint for the air mass travelling below 2 km height and arriving over Athens between 1.5 and 3.0 km a.s.l. at 20:28–21:28 UTC on 22 May 2014. The colors represent the logarithm of the integrated residence time in a grid box in seconds for 6-day integration time.

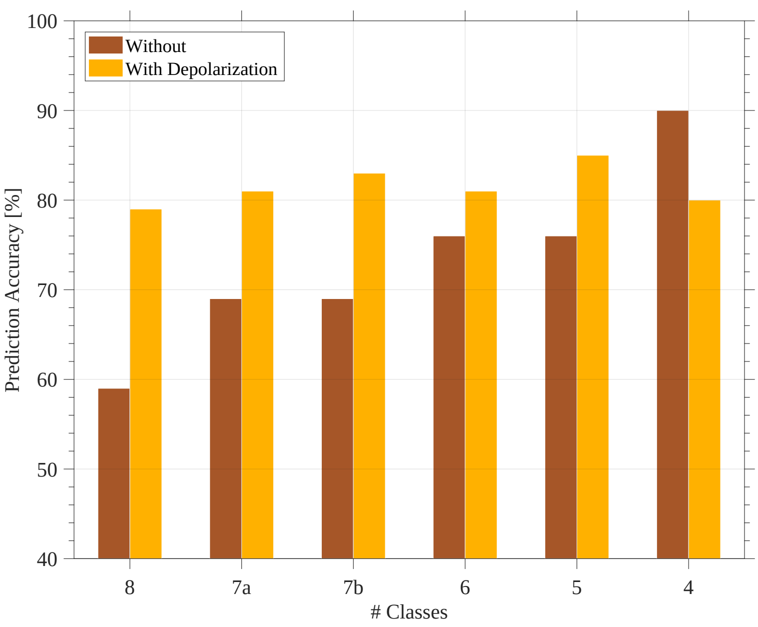

Figure 10. Prediction accuracy for the different aerosol classes and with/without depolarization information.