

Interactive comment on Comparison of two automated aerosol typing methods and their application on an EARLINET station” by Kalliopi Artemis Voudouri et al.

Anonymous Referee #1

As indicated by the title, this paper uses EARLINET 3 α + 2 β lidar measurements to compare the performance of two different aerosol typing algorithms: the Papagiannopoulos (AKA Mahalanobis distance) algorithm and the NATALI algorithm from Nicolae et al.

For the Papagiannopoulos algorithm, aerosol classes are defined using an ensemble of high quality EARLINET measurements that has been manually partitioned by subject matter experts to ensure uniformity of the optical characteristics of all observations within each class. For the Nicolae/NATALI algorithm, aerosol classes are defined using “values obtained from a specially designed aerosol model with simulations from over 50,000 synthetic cases of [different types of] aerosols”. Both classification schemes are trained using supervised learning techniques (although this point is not mentioned in the manuscript).

The lidar measurements used as algorithm inputs for the Papagiannopoulos algorithm are

- a) backscatter-related Ångström Exponent at 355nm and 1064nm (AE 355/1064),*
- b) the lidar ratio at 532 nm (LR 532), and*
- c) the ratio of the lidar ratios (LR 532 /LR 355)*

The lidar measurements used as algorithm inputs by the Nicolae/NATALI algorithm are not specified in this manuscript.

The high resolution variant of NATALI identifies 14 distinct aerosol types. The “low resolution without depolarization” variant used in this study recognizes five aerosol types: dust, smoke, continental polluted, continental, and maritime (page 6, line 15–16). For this study, the smoke and continental polluted types were combined into a single ‘polluted smoke’ type, so that only four types are recognized: dust, polluted smoke, continental, and maritime.

The standard Papagiannopoulos algorithm identifies 8 aerosol types: clean continental, polluted continental, pure dust, mixed dust (dust + maritime), polluted dust (dust + smoke and/or pollution), mixed maritime, smoke, and volcanic. For this study, these 8 types are combined to yield the same four types (page 7, line 1–2) that are recognized by the modified NATALI algorithm: dust (dust + volcanic + mixed dust + polluted dust), polluted smoke (smoke + polluted continental), clean continental, and maritime (mixed maritime).

Both algorithms were applied to the identical set of lidar measurements of 116 different aerosol layers. The NATALI algorithm successfully classified 80 of the 116 layers (~69%). The Papagiannopoulos algorithm successfully classified 114 of the 116 layers (~98%). The subsequent analyses are largely devoted to comparing classification results for the 80 cases in which both algorithms were successful.

The poor performance of the NATALI algorithm when applied to real-world lidar measurements suggests a disparity between the NATALI aerosol models and the EARLINET measurements.

Similar poor performance (39% failure rate) was found in a simulation study in which the authors used the NATALI aerosol model to generate inputs for the Papagiannopoulos algorithm.

Unfortunately, the authors do not investigate the possible causes or sources of these disparities.

In its present form, I cannot recommend this manuscript for publication. Most importantly, this study does not appear to address any specific science questions. Instead, it offers a one-sided comparison between two

retrieval techniques. I say one-sided because the authors' analyses are so heavily focused on those cases where both algorithms are successful. However, they have clearly demonstrated that, for given sets of inputs, both algorithms are subject to large failure rates. The underlying causes of these failures are tantalizingly interesting and, I suspect, likely to be highly informative in developing the next iteration of automated aerosol typing algorithms.

The reviewer is right. In the manuscript we did not explain satisfactorily the main objective of the paper and we did not investigate the possible causes of failure in typing. In the revised version of the paper these comments have been extensively taken into account discussing more these aspects (motivation, agreement and disagreement of the methods). Some text has been added about this and it is reported below into the replies to specific comments. Before addressing the specific comments we would like to give some general remarks about the revised version of the manuscript, which also aim to outline the science questions that are addressed in the paper.

Aerosol typing is an important topic recognized at international level as such: the AEROSAT (International Satellite Aerosol Science Network) initiative established a working group about aerosol types. This is because inferring the aerosol source is an important tool for identification of events, alerting of dangerous situations, definition of mitigation strategies. All these applications are of interest of wide communities. Many remote sensing, near surface and modeling communities have defined ways for inferring/estimating aerosol types and sources and there is the need for defining a common background on this (even nomenclature is not harmonized nowadays) and to compare aerosol type (which is not a numerical value) derived by different methods for improving common understanding and providing a sort of uncertainties on this quantity. In that context in our paper we try to address the following science questions:

- Is there a common understanding for the aerosol types between different typing schemes?
- Can automated methods provide operationally accurate typing when applied in data stored in a climatological database?
- How can we compare different typing methodologies? Which parameters affect the agreement or disagreement between them and the ability to provide typing results.

To address these questions we compare and test in this paper, two independently developed methods based on different approaches. The methods are only briefly described in the manuscript because the respective algorithm papers have already been published in literature. Up to now, there is not a well-established typing scheme and there is actually not an agreed method for evaluating such procedures. The comparison with manual analysis reported in the corresponding algorithm papers and here, however show a promising potential. The idea here is that the convergence of two different methods on the same type is a signature of reliability. This is the reason why emphasis was given to the cases where the two methods agreed. On the other hand, the reviewer is absolutely right: the cases where there was no agreement are extremely important for understanding the limitations of the assumptions behind the two schemes and provide a useful data set to investigate the causes, which will help to potentially improve the typing methods.

This version of the manuscript also contains a number of flaws that make it difficult to understand. For example, a) important terms are never defined (e.g., color ratio);

The reviewer is right. All the intensive optical parameters used in the comparison, are defined properly and all the equations are provided (Section 3) in the revised version of the text.

b) terms that are defined are used inconsistently (e.g., 'Ångström exponent', which is initially defined in terms of backscatter coefficients, is used interchangeably to describe backscatter-related Ångström exponents and extinction-related Ångström exponents);

The text is modified accordingly: "The intensive properties relevant to this study are: the extinction-related Ångström exponent (AE), the backscatter-related Ångström exponent (BAE), the ratio of the backscatter coefficients profiles (color ratios), the lidar ratio (LR), and the ratio of the lidar ratios (RLR). The respective formulas are provided in the following equations, where λ is the wavelength, z is the height, a is the aerosol extinction coefficient, and b is the aerosol backscatter coefficient."

c) the descriptions of the classification algorithms are far too brief (e.g., the description of the NATALI algorithm does not provide the level of detail required to evaluate the sensitivity study that the authors perform); and

A new paragraph with the description of the steps of the Mahalanobis distance-based typing algorithm is added, thus, providing detailed information about the reference dataset and the training of the classifier. Check our answer in Comment 26. Also, a paragraph describing the sensitivity test in NATALI software is added (see Comment 31).

d) the derivation of key results is not well-explained (e.g., the calculation required to produce the numbers reported in figures 4 and 5 are not described).

The revised version is changed accordingly, explaining the derived percentages. In Figure 5 percentages are replaced with numbers. All the numeric results, reported in figures 4 and 5 are now explained in detail in the text. See answer to comment 52 and 54.

Below please find an annotated version of the manuscript the contains several more detailed comments and questions.

We would like to thank the reviewer for his/her fruitful comments that led to the improvement of the manuscript. In the following, answers to comments are reported just below each related comment. When needed, the part of the manuscript we modified or added to the old version, is reported. Additionally, after taking into account the feedback from all the reviewers we proceeded to the following major changes in the revised version of the manuscript.

1) Line 7 page 1 how do you know? what independent measurements or external metrics were used to determine typing success? I hope this topic is explained in detail in the main body of the manuscript. (note: algorithm convergence is not synonymous with successful typing.)

The reviewer is right. We modified the sentence, which now reads: "71 layers out of 110 (percentage of 65%) were typed by both techniques and 56 of the 71 layers (percentage of 79%) were attributed to the same aerosol classes."

2) Line 9 page 1 another note on "successful typing": the fact that both algorithms identify the same type does not mean that both algorithms are correct.

The reviewer is right. We modified the sentence as follows: "The analysis showed that the two algorithms, when applied to real atmospheric conditions, provide typing results that are in agreement (88% for PollutedSmoke, 78% for Dust and 65% for CleanContinental)."

3) Line 10 page 1 how different are the definitions of the types specified by the two algorithms? what degree of difference or similarity should we expect in the typing results based solely on the specifications for the different aerosol classes?

In order to address this comment we added the following paragraph in section 4.4 of the revised version of the manuscript and a summary table (Table 3) which provides mean aerosol optical properties of the reference aerosol types used on the two automated algorithms.

Overall, there were 15 cases (on Thessaloniki dataset) that the two methods provided different typing results. There are seven cases typed as CleanContinental aerosol by NATALI and PollutedSmoke or CleanContinental by EMD, five cases typed as PollutedSmoke by NATALI and Dust by EMD, 2 cases typed as Dust by NATALI and CleanContinental by EMD and one case type as Maritime by NATALI and CleanContinental by EMD. These mismatches are illustrated in Figure 7. In order to understand these differences we highlight some critical issues relevant the type definitions of the two methods, based on Nicolae et al., (2018) and Papagiannopoulos et al., (2018).

CleanContinental: The contribution of the Soot component in the chemical composition of the CleanContinental category, allows higher lidar ratio values at 532nm (52–53sr) in the NATALI scheme. Consequently, layers recognized by NATALI as CleanContinental ones, in the EARLINET Mahalanobis distance-based typing method are attributed as signature for PollutedSmoke or Dust (as it can be seen in the revised Figure 6 -now Figure 7).

Marine: As observations of pure maritime particles are quite scarce within EARLINET and, generally, when these particles are observed, their characteristics are far from pristine, the Maritime category for the EARLINET Mahalanobis distance typing algorithm, corresponds to mixed maritime layers. This is different from the pure maritime category that NATALI identifies. Lower lidar ratio values for 532nm (19–25sr) are defined in the NATALI software for the identification for marine layers, in contrast to the higher ones (16-32sr) allowed in the EARLINET Mahalanobis distance typing algorithm.

Dust: Higher values of Lidar ratio at 532nm are allowed for EARLINET Mahalanobis distance typing algorithm in identifying dust particles (48-62sr), considering all dust-like aerosol types as one category, while NATALI allows values corresponding to more pure cases (44–49). Therefore, a number of Dust recognition are attributed to PollutedSmoke particles (6 cases) or CleanContinental particles (1 case) in the NATALI output.

PollutedSmoke: An almost perfect score was found for the PollutedSmoke category. This can be attributed to the similar reference values attributed by both typing algorithms. Higher values of Lidar

ratio at 532nm are allowed for NATALI (62-92sr - ContinentalPolluted and Smoke), instead of the lower ones (52–89sr) allowed in the EARLINET Mahalanobis distance typing algorithm.

4) Line 16 Page 1 why? what information does aerosol classification provide? is this information critically important all by itself, or solely as input to other applications (or both)? please explain. while AMT is a technology journal, ACP is a science journal. it is therefore essential to clearly explain the scientific importance of aerosol classification.

The reviewer is right. The scientific importance of aerosol classification schemes should be provided in the text. Therefore, we modified the first paragraph of the Introduction, which now reads:

“Aerosol classification is a key parameter in understanding the impact of different aerosol sources on climate, weather systems, and air quality. Given that, aerosols are emitted from various sources, a significant number of aerosol types with different optical and microphysical properties coexist in the atmosphere. As a consequence, considerable uncertainties have been raised in the quantification of aerosols in the radiative transfer calculations and their interactions with clouds, leading the interest of wide communities for improving common understanding related to aerosol types and sources.

Nowadays, aerosol typing is an interesting topic recognized at international level, as inferring the aerosol source is an important tool for identification of events, alerting of dangerous situations, definition of mitigation strategies. The International Satellite Aerosol Science Network (AEROSAT), for example, is an initiative established from different research groups around the world focusing on aerosol retrieval. AEROSAT identifies the need of a common background for aerosol typing and typing procedure’s comparison efforts as one of the main important task to be tackled into the aerosol community (Mona et al., 2015).”

Also, the following sentences are added in the introduction:

“Therefore, the need for defining a common background on aerosol types (even nomenclature is not harmonized nowadays) and for comparing aerosol types (which is not a numerical value) derived by different methods is on the top of interest.”

“This paper can be considered as a step towards the general objective of finding translating rules and way for quantifying differences between typing procedures.”

Also, the following sentences are added in the Sect. 5

“Considering that up to now, there is not a well established method to provide correct results under every conditions, the systematic and full comparison of different methods is the only way to have a better insight about the possibility of inferring the aerosol source from the observed optical properties. Additionally, the modularity of the EARLINET Single Calculus Chain tool offers the possibility to implement aerosol typing procedures for automatically providing the aerosol typing information. A comparison between the two existing typing schemes that make use of the aerosol optical properties could be rather useful. This will allow also from one side the optimization of aerosol

property models and from the other side the definition of the aerosol types used and the selection of the reference dataset.”

5) Line 18 Pag. 1 do you really mean "magnitudes" (which can be directly measured) rather than "mechanism" (which must be inferred)

The text has been modified to:

“Thus, a number of typing schemes have been applied to classify aerosols, based on the synergistic information of the backscattering and extinction coefficients and the polarization state of the received light at different wavelengths (e.g, Ansmann and Müller, 2005, Groß 2013).”

6) Line 19 Page. 1 according to Hamill et al. (2016), “Our aerosol characterization is based on the AERONET retrieved quantities, therefore it does not include low optical depth values.” somewhere in this manuscript, I hope the authors will discuss the range of optical depths over which each method can reliably ascertain aerosol types.

The reviewer is right. When examining the ability of the methods to provide a classification result for a certain aerosol type or mixture the uncertainties of the measurements have to be taken into account. As large measurement uncertainties prevent a correct aerosol type separation, high quality measurements are mandatory. Some criteria apply for the layers identification (Belegante et al., 2014). In particular, the relative error of each intensive optical parameter calculated as mean within the layer should be lower than 50 % in order to be considered for the typing procedures. Additionally, the range of the parameters to ensure reliable retrievals of aerosol types, are discussed in terms of the values of the extensive or intensive parameters that were used in the typing modules. So, a table with the acceptable limits for the average intensive optical parameters was added in the text. In addition, we examined the ability of the both methods to provide typing results, relative to the derived layer AOD values. This study showed that the mean layer AOD of the typed layers has similar value for both typing techniques (0.1076 ± 0.0898 (SD) for Natali and 0.11166 ± 0.0928 (SD) for the EARLINET Mahalanobis distance typing algorithm (EMD). However, the EARLINET Mahalanobis distance algorithm seems more sensitive to lower values (not typed cases show a mean layer AOD of 0.0298 ± 0.0578 for EMD, while 0.06899 ± 0.0606 for NATALI). This information has been included in the revised manuscript.

7) Line 5 Page. 2 In addition to these examples, please provide definitions for both intensive and extensive the key difference is that extensive properties vary with aerosol concentration (or amount or loading), whereas intensive properties are independent of concentration and instead vary only according to aerosol composition (i.e., type). it should also be noted even intensive properties might not be sufficient to guarantee accurate typing, as some aerosol types (e.g., biomass burning and industrial pollution) have very similar intensive properties but very different sources and/or generating mechanisms. these definitions should be provided the first time the terms are used.

We thank the reviewer for this comment. The text included is given below:

“The extensive properties depend on the aerosol concentration, whilst intensive ones are type sensitive and provide separate classification for each detected layer (e.g., Müller et al., 2007; Ansmann et al., 2011; Tesche et al., 2011; Burton et al., 2012; Pappalardo et al., 2013; Groß et al., 2013; Amiridis et al., 2015; Giannakaki et al., 2015; and Baars et al., 2016). Nevertheless, even intensive properties might not be sufficient to guarantee accurate typing, as some aerosol types (e.g., biomass burning and industrial pollution) have very similar intensive properties but are attributed to different sources and generating mechanisms.”

8) Line 10-11 Pag. 2 also CATS; see Yorks et al., 2016: CATS Algorithm Theoretical Basis Document, Level 1 and Level 2 Data Products, https://cats.gsfc.nasa.gov/media/docs/CATS_ATBD.pdf

references? (e.g., Burton et al., 2012)

references? (e.g., Papagiannopoulos et al., 2018)

Thank you for this comment. The appropriate references have been added.

“Aerosol typing schemes have been developed for high-resolution lidar measurements of space-borne lidars (e.g., CALIPSO, Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation, Omar et al. 2009; EarthCARE, Illingworth et al. 2015, CATS, Yorks et al., 2016), airborne - High Spectral Resolution Lidar -HSRL- measurements (e.g., Burton et al., 2012; Groß et al., 2013) and multiwavelength Raman measurements (e.g., Nicolae et al., 2018; Papagiannopoulos et al., 2018).”

9) Line 26 page 2 I'm impressed by this very extensive and complete list of references

We thank the reviewer for acknowledging the efforts of EARLINET.

10) Line 4 Page 3 this statement needs references. off the top of my head, I would expect the extinction Angstrom exponents of volcanic ash and dust might be quite different.

Yes, it is true that volcanic ash and desert dust can have different sizes depending on the transport pattern and aging processes. Indeed, many differences can be observed in the optical properties even among volcanic particles depending on the specific volcano and the eruption type and among desert dust particles depending on the source region. What we would like to underline here is that volcanic and desert dust particles are both characterized by the asphericity of the particles, observable and measured through the linear depolarization ratio (while typically the other aerosol types have small values of this parameter), the other intensive optical properties could be similar in ranges e.g. Angstrom exponents and lidar ratio (see for example fig. 5 in Nicolae et al., 2018; fig. 7 in Papagiannopoulos et al., 2018 and the variety of volcanic values in Mona et al., 2012). However, because the source region for these cases is well defined the use of transport models could allow for a more precise attribution of the aerosol source. EARLINET-based studies for the Eyjafjallajökull eruption have indicated similar Ångström exponent values to desert dust. Specifically, Papayannis et al., (2012) found mean Angstrom exponent values of 0.57 ± 0.26 . Furthermore, Sicard et al. (2012) and Ansmann

et al. (2010) calculated a mean Ångström exponent of 0.68 ± 0.63 and 0.03 ± 0.4 for volcanic ash respectively. Following the suggestion of the reviewer, references have been added:

“Furthermore, the combined use of lidar observations and transport model simulations permit the discrimination of desert dust and volcanic ash particles that typically, have the same optical characteristics (see for example fig. 5 in Nicolae et al., 2018; fig. 7 in Papagiannopoulos et al., 2018 and the variety of volcanic values in Mona et al., 2012). Simultaneous observations of desert dust and ash particles were made during the Eyjafjallajökull volcanic eruption in 2010 and the methodology for the type discrimination was presented by Papayannis et al. (2012); Mona et al. (2012); and Pappalardo et al. (2013).”

11) Line 8 Page 3 what science questions are being addressed by this comparison?

As stated also in our general remarks the aim of this comparison is to check the performance of the two techniques to a specific environment, focusing on the reasons that they agree or disagree and the limitation of each in the aerosol typing. Check the performance of each different reference dataset is really important for defining a common understanding on aerosol types (even nomenclature is not harmonized nowadays) and for evaluation strategies to compare aerosol types (which is not a numerical value) derived by different methods. In that context the applied thresholds, differently introduced in each algorithm, are a critical factor in the identification rate of the two algorithms. The latter is discussed in more detail in the revised manuscript.

The next paragraph is added to the manuscript and we think that pinpoints the importance of this work:

“Aerosol typing schemes have been discussed to be implemented in the next version of the EARLINET Single Calculus Chain tool (SCC; D’Amico et al., 2015). SCC was developed for the analysis of the data of different lidar systems in an automated, unsupervised way and currently SCCv5.0.11 delivers profiles of optical aerosol properties. Future work is dedicated to the implementation of new features like profiles of intensive optical properties, detection of aerosol layer geometrical properties and calculation of the intensive optical properties per detected layer, which will further allow the classification of the observed layers into aerosol types. Therefore, the need for defining a common background on aerosol types (even nomenclature is not harmonized nowadays) and for comparing aerosol types (which is not a numerical value) derived by different methods is on the top of interest.

Given the aforementioned status, a first attempt of comparing and evaluating two classification tools developed within EARLINET, that provide near-real-time aerosol typing information for the lidar profiles of Thessaloniki, is presented. Our aim is: (i) to check the performance of both supervised learning techniques in their low resolution mode, which is the case for the majority of the available measurements within EARLINET, (ii) to highlight the strong and weak aspects of both algorithms when applied to lidar data from a station where, typically, variable mixtures of aerosols are present and (iii) to investigate the reasons of typing agreement and disagreement with respect to the uncertainties and the threshold criteria applied. This paper can be considered as a step towards the general objective of finding translating rules and way for quantifying differences between typing procedures.”

12) Line 23 Page 3 unclear; were particulate depolarization ratios previously available? or have the detectors and optics only recently been added, and hence the depolarization measurements are not yet part of the standard data stream?

The two depolarization channels have been added to measure the cross and parallel polarized signal at 532 nm, but due to issues, related to the polarization purity of emitted laser beam and the performance of the telescope lenses, the particle depolarization ratio was not available. A new telescope was installed at the end of 2017, thus for the period under study no depolarization measurements are available. Indeed, these issues could also partially affect the total 532 nm elastic backscatter measurements

So, according to the above, the text has been modified to:

“These two channels have been added to measure the cross and parallel polarized signal at 532 nm, but due to technical issues the particle depolarization ratio is available after August 2018.”

13) Line 2 Page 4 why is depolarization enclosed in parentheses? if depolarization ratios are optional but not actually required, they shouldn't be cited as requirements.

The reviewer is right. The text has been modified to:

“The two automatic aerosol typing methods require only lidar data with $3\beta + 2\alpha$ (3 backscatter and 2 extinction coefficient profiles) configuration without any use of ancillary external information.”

14) Line 2 Page 4 this seems like an odd choice to me. why not use the extinction Ångström exponents instead? surely the extant literature contains much, much more information about extinction Ångström exponents (which can be obtained by any number of sensors and are of interest in a wide variety of research applications) than about the ratio of lidar ratios (which are only of direct interest to lidar practitioners).

The ratio of lidar ratios (RLR) is a quantity that provides information about the aerosol size, and is preferred due to the fact that, the extinction coefficient, is associated with larger uncertainty. This is attributed mainly to the derivative in the inversion algorithm (e.g, Ansmann et al., 2002; Freudenthaler et al., 2009). Also, the Ratio of Lidar Ratios was chosen as a classification parameter, after the performed sensitivity analysis from Papagiannopoulos et al., (2018), to identify which classifying properties provide the adequate information to better predict the correct aerosol class. Two statistical parameters were used in the analysis: the total and the partial Wilks' lambda (3; Wilks, 1963) that are widely used, e.g, Burton et al. (2012) and Russell et al. (2014).

RLR quantity is not only available by lidar systems, but can also be obtained by the Sunphotometers of the Aerosol Robotic Network (AERONET), by dividing the provided lidar ratios, (Shin et. al., 2018).

In the revised paper the introduction of Section 3 reports all the definition and meaning of the different parameters and then it is shortly recalled the work done in Papagiannopoulos et al., 2018 for selection of how many and which parameters are the most suitable from a mathematical point of view for the typing purposes.

“The intensive properties relevant to this study are: the extinction-related Ångström exponent (AE), the backscatter-related Ångström exponent (BAE), the ratio of the backscatter coefficients profiles (color ratios - CI), the lidar ratio (LR), and the ratio of the lidar ratios (RLR). The respective formulas are provided in the following equations, where λ is the wavelength, z is the height, a is the aerosol extinction coefficient, and b is the aerosol backscatter coefficient.”

15) Line 10 page 4 please provide the equation

All the equations used for the typing are included in the text (in Sect. 3 “Aerosol Typing Methods”).

16) Line 10 page 4 wouldn't extinction Ångström exponents be a more accurate and reliable size indicator? (certainly extinction Ångström exponents have been investigated much more thoroughly in previous publications than have backscatter Ångström exponents).

As extinction coefficient, is associated with larger uncertainty, the backscatter-related Ångström exponent is preferred. There is a number of studies that have used in their aerosol classification this parameter, e.g., Müller et al. (2007). Besides, the backscatter related Ångström exponent has been extensively used in the climatological analysis made for Thessaloniki, EARLINET station, by Amiridis et al., (2009), Gianakkaki et al. (2010), and Siomos et. (2018).

16) Line 20 page 4 what are color ratios? please define this term.

The term is now described in the revised manuscript. The text has been modified to:

“The extinction-related Ångström Exponent at 550nm and 350nm (AE 350/550), the lidar ratio at 350 nm (LR 350), the lidar ratio at 550 (LR 550), and the ratio of the backscatter coefficients profiles (color ratios) CR 350/550 and CR 550/1000 , are used for the aerosol classification.”

18) Line 23 page 4 this seems misleading to me. according to Nicolae et al. (2018), the required data are “backscatter coefficient (β) profiles at 1064, 532 and 355 nm, extinction coefficient (α) profiles at 532 and 355 nm, and, optionally, linear particle depolarization (δ) profile at 532 nm.”

The reviewer is right and indeed this was not clear in the manuscript. The next few lines are included in the revised manuscript:

“The input module requires optical properties profiles as those measured by EARLINET stations, namely the aerosol extinction coefficient and the aerosol backscatter coefficient profile. Optionally, the linear particle depolarization profile at 532 nm can be provided, so as to allow a better classification and to increase the number of classified aerosols types.”

Also, the following paragraph is modified:

Depending on the availability of the particle linear depolarization ratio and the quality of the provided lidar profiles, the derived typing can be either of high resolution (AH), or low resolution with depolarization (AL) or low resolution without depolarization (BL). Pure aerosols categories, and even mixtures of three aerosols types can be obtained from the NATALI algorithm. In the high resolution typing, 14 aerosol types can be distinguished (i.e., Continental, ContinentalPolluted, Dust,

Maritime/CC, Smoke, Volcanic, Coastal, CoastalPolluted, ContinentalDust, ContinentalSmoke, DustPolluted, MaritimeMineral, MixedDust and MixedSmoke) when the quality of the provided optical products is high enough. In the low resolution typing (AL), 6 predominant aerosol types can be provided but with high uncertainty (i.e, Continental, ContinentalPolluted, Smoke, Dust, Maritime and Volcanic). The low resolution typing (BL) provides 5 predominant aerosol types (Dust, ContinentalPolluted, Smoke, Continental, Maritime) either pure or mixed, when the depolarization information is not provided. Finally, the output module provides the intensive optical parameters within each layer along with their mean value and the corresponding uncertainty.

19) Line 26 page 4 I'm not sure I understand what's meant here. are the authors computing weighted means, where the weights are the SNRs at each data point? if so, won't in-layer attenuations and the attenuation of the signals between successive layers adversely affect the SNR of the points at higher altitudes, thereby biasing their contributions to the mean values?

No. We mean that we calculate mean properties only for layers where the SNR is greater than 5. The text is changed accordingly.

20) Line 30 page 4 either "most likely" or "most probable", but not "most likely probable"

This was a typo and it is now fixed.

"The identification of the most probable aerosol type is then made through a voting procedure using the results from the three ANNs interrogated."

21) Line 32 page 4 since this is a "specially designed aerosol model", why does it use wavelengths that are different from the measured wavelengths?

The text has been modified accordingly:

"The calculations of the optical properties for each layer are used in the typing module and compared to values obtained from a specially designed aerosol model with simulations from over 50000 synthetic cases of aerosols. Specifically, the synthetic database was developed using the aerosol model built for 350, 550, and 1000 nm sounding wavelengths, based on the 61 wavelengths of OPAC (Optical Properties of Aerosols and Clouds) software package (Hess et al., 1998), for which the microphysical characteristics of the aerosols are available from GADS. These wavelengths are then re-scaled to the usual lidar wavelengths (i.e., 355, 532 and 1064 nm) using an average Ångström exponent equal to one. This was considered a valid assumption for all aerosol types, taking into account the small difference between the lidar and the model wavelengths. A comprehensive description of the developed aerosol model can be found in Nicolae et al., (2018)."

22) Line 32 page 4 need to define color ratio

The term is now described in the revised manuscript (also the equation is provided in Section 3).

“The extinction-related Ångström Exponent at 550nm and 350nm (AE 550/350), the lidar ratio at 350 nm (LR 350), the lidar ratio at 550 (LR 550), and the ratio of the backscatter coefficients profiles (color ratios) CR 550/350 and CR 1000/550 , are used for the aerosol classification.”

23) Line 32 page 4 presumably the 350 nm, 550 nm and 1000 nm wavelengths are specified by the model. please state definitively which lidar measured parameters are used in the configuration of the NATALI algorithm used in this paper.

The next few lines are included in the revised manuscript:

“The synthetic database was developed using the aerosol model built for 350, 550, and 1000 nm sounding wavelengths, based on the 61 wavelengths of OPAC (Optical Properties of Aerosols and Clouds) software package (Hess et al., 1998), for which the microphysical characteristics of the aerosols are available from GADS (Global Aerosol Date Set). These wavelengths are then re-scaled to the usual lidar wavelengths (i.e., 355, 532 and 1064 nm), that are used as input to NATALI, using an average Ångström exponent equal to one.”

24) Line 9 page 5 so all NATALI results used in this comparison study will be BL only, yes?

Yes, the comparison is made for the low resolution without depolarization (BL). A Table is added to provide the information of the resolution and the aerosol types that constitute the classes investigated in this study.

25) Line 9 page 5 do you mean "along with" here?

The reviewer is right. The sentence now becomes:

“Finally, the output module provides the intensive optical parameters within each layer along with their mean value and the corresponding uncertainty.”

26) Line 25 page 5 seems like you're omitting an important step here. see section 3.2 in Papagiannopoulos et al., 2018. before you can start classifying previously unclassified data, don't you first have to train your classifier to recognize the manually specified classes in the reference dataset established in step 1?

The reviewer is right. The text has been modified to:

“The classification of each aerosol layer is made by calculating the distance of an observation from the defined reference classes and by attributing each observation to a specific class based on the minimum distance. Given that the overall predictive accuracy of a typing algorithm, depends on the defined reference classes, the choice of the appropriate reference dataset consists a crucial parameter. So, a well-characterized EARLINET dataset of observations from 2008 to 2010 was used to define the aerosol classes (Pappalardo et al. 2013; Papagiannopoulos et al. 2016a, and Schwarz 2016). As this dataset did not include all the typical aerosol components presented over Europe, the observations were enhanced with additional ones (already published in literature), contributing to a total of 69 layers (Pappalardo et al., 2013; Papagiannopoulos et al., 2016a), corresponding to the

following types: CleanContinental (9 cases), PollutedContinental (16 cases), Marine (8 cases), Dust (9 cases), MixedDust (10 cases), PollutedDust (5 cases), Smoke (7 cases), Volcanic ash (5 cases).

In a next step, a sensitivity analysis was performed in order to identify the classifying properties that provide the most adequate information to the better prediction of the correct aerosol class. The intensive properties that held the most weight among others in the classification appeared to be the backscatter-related Ångström Exponent at 355nm and 1064nm (BAE 355/1064), the lidar ratio at 532 nm (LR 532), and the ratio of the lidar ratios (LR 532 /LR 355). Then, the initial reference dataset was split into a training and a validation dataset, in order to evaluate the predictive accuracy of the automatic method. Therefore, the training dataset was used to calculate the classification functions and then each observation was validated using the training dataset. Additionally, the effect of the depolarization information to the training phase was also investigated. To complement the reference dataset in this context, general literature values for particle linear depolarization ratio at 532 nm were used. Generally, the depolarization ratio is a really important parameter that can strengthen the typing procedure in the presence of aspherical layers, but Papagiannopoulos et al. (2018) showed that for the 5 and 4 classes, the particle linear depolarization ratio became less important (in this case the highest weight in the classification corresponds to the lidar ratio at 532 nm)."

27) Line 25-30 page 5 the description here does not provide enough detail for me to grasp the essentials of the algorithm. please expand this section.

See above, the answer to Line 25 page 5.

28) Line 31 page 5 what is the physical meaning of this "normalized probability" and how does one compute it? this term should be clearly defined in the manuscript.

also, I'm assuming that the authors mean that "normalized probabilities" are computed for all candidates, and that for a successful classification to occur, the largest "normalized probability" must exceed 0.5. please add text to clarify.

The Mahalanobis distance quantifies the distance between a point (i.e., the unclassified measurement) and a distribution (i.e., the aerosol class). This distance is estimated for each of the available number of classes and, thus, it is independent. Besides, the Mahalanobis distance gives an estimate of the probability for each aerosol class. This probability has to be normalized so as to provide an estimate of a relative probability for each class. The normalized probability ensures that similar class distances will not be classified. The following equation is used to derive the normalized probability:

$$\frac{1}{d^2} \bigg/ \sum \left(\frac{1}{d^2} \right),$$

where d stands for the distance. More details can be found at Burton et al. (2012). A

proper explanation is included in the revised manuscript.

29) Line 6 page 6 classifications showed a high success rate

The text has been modified to:

“Dust classifications showed a high success rate, whilst the aerosol types that performed worse were the smoke and polluted continental aerosol.”

30) Line 19 page 6 should this be 'minimum layer thickness'?

The reviewer is right. This should be the minimum layer thickness.

31) Line 19 page 6 the authors need to provide the following information before readers (and reviewers) can evaluate the effectiveness and completeness of their sensitivity analysis.

(a) where and how is the layer thickness parameter used?

(b) what is being smoothed and where (and how often) in the analysis is this smoothing applied?

(c) what quantities are the minimum confidence and minimum agreement ratio applied to?

(a) Layer thickness parameter is defined by the user in the graphical interface of NATALI module. The parameter is defined in meters, so as NATALI module to reject layer sub-structures thinner than the selected value.

(b) Smoothing is applied to the derivative of the 1064nm signal for the definition of the boundary layers. The inflexion points of the second derivative of the profile data, computed with the Savitzky-Golay filter, give the boundaries of the layers. The window size of the cubic Savitzky-Golay filter, which can be modified by the user, has a default value of 700 m. The filter was applied twice to obtain the second derivative. A signal-to-noise ratio filter is applied at this point, making sure the said ratio is at least 5. The layer boundaries are moved towards the median height until the SNR criteria is met; if the criteria cannot be satisfied with a layer height greater than e.g., 300m, the layer is discarded. Gross or fine structure of the aerosol layers is revealed by a higher or lower value of the adjustable smoothing parameter (FINESSE).

(c) The identification of the most probable aerosol type is made through a voting procedure, using the results from the three ANNs interrogated. Over 50000 aerosol synthetic data have been used to train the ANN and identify the better ANNs to classify the aerosols type from multiwavelength lidar data. The capability of ANNs to resolve the overlapping clusters of the intensive optical parameters is used on NATALI algorithm. The answer selected is based on a statistical approach. The selected types of ANNs classify the aerosols based on the response with high i) confidence (i.e. the probability of having one of the aerosol types) and ii) stability over the uncertainty range (i.e., the percentage of agreement for values between error limits). Therefore, answers with low confidence are filtered out and NATALI returns the 'Unknown' type. In this study, we select initially the confidence level for the output retrievals higher than 0.9 (minimum accepted confidence) and the minimum agreement threshold as default (i.e., 0.25).

The text has been modified accordingly in the revised manuscript.

32) Line 20 page 6 I'm afraid I have no idea what this sentence is supposed to tell me. this is the first mention of smoothing in the manuscript. what is being smoothed, and why are you smoothing it? 700 number of values generated between value \pm uncertainty" sounds like the description of a simulation study. how does this relate to smoothing?

The reviewer is right. This is stated wrongly and not related with the smoothing. The selected (from the user) number of values (10-50) generated between [mean intensive optical parameter – uncertainty] and [mean intensive optical parameter + uncertainty] correspond to finesse parameter. Optical parameters calculated from lidar measurements are reported in the EARLINET database with the mean value and the associated uncertainty, but the optical parameters calculated from synthetic data do not “carry” this uncertainty. For this reason, a fixed relative error was considered, which was multiplied with the mean value to obtain the absolute error (uncertainty). For the actual retrieval of the aerosol type, any value between within mean value +/-uncertainty was possible; therefore the algorithm was applied for all these values with a certain step (i.e. the finesse). The output is a “bundle” of possible aerosol types, with a dimension equal to the finesse.

The text in the revised manuscript is changed accordingly.

33) Line 20 page 6 I'm bewildered by the repeated use of the word "value". do you mean the number of in-layer values lying between the characteristic mean \pm uncertainty for each parameter? how does this relate to smoothing?

Indeed, values correspond to the number of in-layer values lying between the characteristic mean parameter \pm uncertainty.

34) Line 33-34 page 6 these two statements contradict one another. please clarify: does the NATALI algorithm identify 5 types or 4 types?

The reviewer is right. It was not stated clearly in the text. NATALI algorithm output in the low resolution module identifies 5 types. We decided to merge the two categories -PollutedContinental and Smoke- to a more generic one, the category PollutedSmoke. The selection of four aerosol classes stems from the difficulty to distinguish types of aerosols that reflect the same aerosol characteristics. Considering the aforementioned typing merging, the EARLINET Mahalanobis distance-based typing algorithm was set to classify observations into 4 aerosol classes. This information is summarized and made more clear in Table 2 of the revised manuscript.

So, the text has been modified accordingly:

“In what follows, we merged the output types from NATALI that tend to reflect the same aerosol characteristics, and hence we evaluate the corresponding effects on the prediction rate of the algorithms. Thus, the smoke and the polluted continental categories were grouped into the more generic type of small particles with high lidar ratio values. The selection of four main aerosol classes stems from the availability of intensive properties, the difficulty in deriving a confident classification

without particle linear depolarization ratio and the difficulty in discriminating polluted continental and smoke particles that reveal the same type characteristics. Regardless, the aerosol classes describe the major aerosol components.”

And later on the methodology part, the following change is made:

“The identified layer boundaries from NATALI are used as input in the EARLINET Mahalanobis distance-based typing algorithm. Considering the aforementioned typing merging, the EARLINET Mahalanobis distance-based typing algorithm was set to classify observations into 4 aerosol classes: CleanContinental, Dust, Maritime and PollutedSmoke.”

35) Line 1 page 7 please provide your rationale for combining types as you've done. Line 1 page 5 again, please provide your rationale for combining these two types. also, to avoid confusing your readers, this information should be given earlier (e.g., around line 15 on the previous page)

We merged types that tend to reflect the same aerosol characteristics; the smoke and the polluted continental category were grouped into the more generic type of small particles with high lidar ratio values. Their intensive properties tend to be overlapped as can be seen in Groß et al. (2013). See also, answer 34. Also, in the low resolution typing all the dust like particles are considered as one category, given that the depolarization ration is not available to distinguish between pure dust and mixed/polluted dust particles. Also, a Table is added to provide all the necessary information.

36) please provide a reference. I know of references supporting the assertion of low extinction Ångström exponents for dust (i.e., close to zero, as in Freudenthaler et al., 2008: doi:10.1111/j.1600-0889.2008.00396.x), but I can't immediately recall any for backscatter Ångström exponents.

Müller et al. (2007) provided an analysis for backscatter Ångström exponents, reported values of 0.2 ± 0.2 for desert dust events. The following sentence was added:

“Also, Müller et al., (2007) reported values of mean backscatter-related Ångström exponent BAE 355/1064 of 0.2 ± 0.2 for desert dust events.”

37) I note that table 4 in Siomos et al. 2018 reports backscatter Ångström exponents for dust mixes in the neighborhood of ~1. a "low Ångström exponent" would be ~0.

The reviewer is right. These values reported by Siomos et al. (2018) correspond to mixed dust cases (and not pure one) typically observed over the urban Thessaloniki station.

38) Line 13 page 7 see my previous comment. table 4 in Siomos et al. 2018 reports backscatter Ångström exponents for dust mixes of ~1. values of ~1 cannot simultaneously be characterized as both "low" and "relatively high"

The reviewer is right and the text is modified to:

“Generally, continental particles (i.e., small with medium lidar ratio values) present low lidar ratio values, (i.e., 20–40 sr) and values of Ångström exponents around 1.0–2.5.”

39) Line 19 page 7 were the same criteria used in the automated typing also used in the manual typing? what (if any) additional information was used in the manual typing exercise that was not available in the automatic typing schemes?

The so-called manual aerosol typing can be described by a three-step procedure (Papagiannopoulos et al., 2016). First, the aerosols and clouds are identified in the lidar data. Second, the boundaries of the aerosol layer are defined. Third, the aerosol layer is typed by means of investigation of intensive optical properties, model outputs (e.g., HYSPLIT simulations), and ancillary instrument (e.g., AERONET sunphotometer data, MODIS FIRMS data). Both automatic methods make use of lidar intensive properties and predict the aerosol type with no further external information. The lidar measurements of intensive properties are shown to vary with aerosol type and are able to provide an objective aerosol characterization.

40) Line 4 page 8 the authors need to do a much better job of defining terms and using them consistently. what's the difference between AE355-532 and BAE355-532? what's the difference between AE355-532 and AE355/532? on page 4, line 10, AE is defined as the backscatter-related aerosol Ångström exponent. at this point in the manuscript, no definition of BAE has been given. scanning ahead to page 8, line 19, I see that BAE is also defined as the backscatter-related aerosol Ångström exponent.

The reviewer is right. The definitions of the backscatter-related aerosol Ångström exponent and the extinction-related aerosol Ångström exponent were not provided clearly in the text. In the revised version, though, everything is stated properly and all equations are provided in Section 3.

The text has been modified to:

“The intensive properties relevant to this study are: the extinction-related Ångström exponent (AE), the backscatter-related Ångström exponent (BAE), the ratio of the backscatter coefficients profiles (color ratios), the lidar ratio (LR), and the ratio of the lidar ratios (RLR). The respective formulas are provided in the following equations, where λ is the wavelength, z is the height, a is the aerosol extinction coefficient, and b is the aerosol backscatter coefficient.”

41) Line 4 page 8 the Papagiannopoulos algorithm used the 355-to-1064 nm backscatter-related Ångström Exponent. why are those values not given here? why are the 532-to-1064 nm backscatter-related Ångström Exponent given? this is the first time they're mentioned in the manuscript, and it's not apparent that they're used in any aspect of either classification scheme.

The reviewer is right. The 355-to-1064 nm backscatter-related Ångström Exponent is provided in the plots in the revised version. The 532-to-1064 nm backscatter-related Ångström Exponent was provided in the plots, because this was one of the optical parameters considered for typing in NATALI.

42) Line 8 page 8 is this a "normalized" probability? The probabilities shown in the figure for the different types do not appear to add up to one. why not? please add text to explain.

The probability reported here is not normalized, but the chi-square probability for each class. This probability is assigned to each distance, it is independent, and shouldn't add up to one. See previous comment (n 28).

The following paragraph is added:

“The output of the algorithm is the calculated Mahalanobis distance of the layer from each aerosol class and the calculated chi-square probability. For improving the reliability of the output, two screening criteria are applied to the minimum distance following the procedure of Burton et al. (2012). In our study, we use 3 classifying parameters (i.e., the type dependent intensive properties) and the minimum accepted distance for a measurement to be labelled is 4. When the distance is higher than the defined threshold, which means that no similarity with the aerosol classes of the reference datasets is found, the observation is not typed. Secondly, an extra screening criterion is applied to almost equal values for more than one aerosol class. The normalized probability (this probability is based on the calculated distances) of the aerosol class needs to be higher than 50%.”

Also, the following sentence is added to Section 4.1 and 4.2.

“The plotted probabilities are assigned to each calculated distance and are different from the normalized ones that used as a screening criteria (and add up to 100).”

Similar comment is made in the captions of Figures 1, 2: ‘EARLINET Mahalanobis distance-based typing algorithm probability for each detected layer (different from the normalized one)’

43) Line 18 page 8 these look like very representative value for the extinction Ångström exponent of smoke (though layers B and C are slightly low). perhaps it's worth adding a reference that established a connection to previously reported values?

Tesche et al. (2011) reported values for the extinction Ångström exponent of smoke 0.9 ± 0.26 . The following sentence is added:

“These values are consistent with previous reported ones, e.g., Tesche et al. (2011a) reported values of the extinction Ångström exponent for smoke of 0.9 ± 0.26 .”

44) Line 20 page 8 again, why report BAE532-1064 when the Papagiannopoulos algorithm uses BAE532-1064? as a referee (and as a reader!) I'm looking for correspondence between the measurements and the characteristic attributes of the aerosol types defined by each algorithm.

The reviewer is right. The 355-to-1064 nm backscatter-related Ångström Exponent was provided in the plots in the revised version. The 355-to-1064 nm backscatter-related Ångström Exponent was provided, because this parameter is one of the parameters taken into account for NATALI typing.

45) Line 30 page 8 so, the EARLINET-Mahalanobis algorithm achieves an almost perfect score (114/116), whereas NATALI fails ~30% of the time? that alone is a very notable performance difference between the two methods. where is the analysis that explains this stark difference? in particular, how do the “manually typed” results compare to the EARLINET-Mahalanobis algorithm classifications for the 34 layers for which NATALI failed but the EARLINET-Mahalanobis algorithm succeeded?

Indeed, the two methods are different. This is general problem of the various aerosol typing algorithms and as already discussed in our general remarks, the international initiative AEROSAT identifies the need of a common background and comparison efforts as one of the main important task to be tackled into the aerosol community and a working group was set up for this purpose. One of the coauthor of this paper is leading this Working group and this paper can be considered as a step towards the general objective of that WG for finding translating rules and ways for quantifying differences between typing procedures. In particular, the Mahalanobis distance typing is designed in such a way that it does not take into account the error on the measurements. It considers the mean measured values as true and when the distance of the measurement from the clusters defined through the reference datasets is large, with respect to a defined threshold, the observation is not typed at all. Comparing methods which are too similar indeed would not add any information to the scene, what is instead interesting is understanding where the 2 methods converge and where they see differences. There is not nowadays one established method which provides accurate results under every condition, so the systematic comparison of different methods is the only way to have a better insight about the possibility of inferring the aerosol source from the observed optical properties.

However it should be pointed out, that the identification rate of the two algorithms can be changed regarding the needs of the user and the selection of the appropriate threshold criteria.

With the used criteria in the NATALI software (with a confidence for the output retrieval higher than 0.9), the algorithm returns only the retrievals that pass a number of quality checks, so as to make sure (as possible) that the output is trustworthy. As such, the percentages of ‘unknowns’ are not because NATALI cannot identify the aerosol type, but due to large uncertainties of the optical parameters and the calculated uncertainties. However, these ‘unknowns’ layers were typed (89% and 11% stayed untyped), when the minimum acceptance confidence were chosen differently (with a confidence higher than 0.7). Specifically, from the 28 ‘unknowns’ layers, 26 were typed when the minimum acceptance confidence is chosen lower, 2 remained ‘unknowns’. From these 26 cases 11 (40%) give the same aerosol type with EARLINET Mahalanobis distance typing algorithm. A new plot is added in the revised version of the manuscript, explaining each classification.

From the mismatches: 7 cases were attributed as Dust from EARLINET Mahalanobis distance typing algorithm (Hysplit back-trajectories don’t agree) and identified from NATALI either as Continental or Smoke. Higher values of Lidar ratio at 532nm are allowed for EARLINET Mahalanobis distance typing algorithm in identifying dust particles (48-62sr), considering all dust-like aerosol types as one category. Therefore, a number of Dust recognition (from EMD) are attributed to Smoke particles or CleanContinental particles in the NATALI output. CleanContinental(NATALI) and Maritime (EARLINET

Mahalanobis distance typing algorithm) don't agree most of the times. Also, a Dust case (typed by NATALI) is attributed as Smoke from the EARLINET Mahalanobis distance typing algorithm (back-trajectories indicate air masses from fires spots).

Nevertheless, we choose to show the comparison, holding the retrievals that pass a higher percentage of quality checks.

Also, when the EARLINET Mahalanobis distance typing algorithm is forced to categorize a layer, only when the normalized probability is above 60% (instead of 50% that used in the comparison), the number of untyped layers increased (20 untyped cases, 81% classification of the total number of cases). These layers are the same that failed to be typed with the higher percentage of quality checks in NATALI. Nevertheless, we show the comparison for the criteria chosen also in the study of Papagiannopoulos et al., (2018).

Once again, we conclude that the identification rate of the two algorithms can be changed regarding the needs of the user and the selection of the appropriate threshold criteria.

Therefore the following paragraph has been added in the revised manuscript/

“NATALI classified 82 of the total 110 layers (75% of the total number of cases) and 28 were flagged as Unknown layers. These missed cases are based mostly on retrievals with larger uncertainties of the optical parameters and the calculated uncertainties, that not allowed ANNs to return any aerosol type. However, these ‘unknowns’ layers were typed (93% and 7% stayed untyped), when the minimum acceptance confidence was chosen lower (i.e., confidence higher than 0.7). Nevertheless, we choose to show the comparison, holding the retrievals that pass a higher percentage of quality checks. The EARLINET Mahalanobis distance-based typing algorithm classified 98 layers (89%) and 12 layers were not assigned to any cluster (the Mahalanobis distance values larger than 4), showing a higher identification rate. However, when we set the second of the screening criteria (i.e., the normalized probability of the aerosol class) to be higher than 50%, the number of untyped layers increased. We conclude that, the identification rate can be changed regarding the needs of the user and the selection of the appropriate threshold criteria.”

46) Line 5 page 9 the dust cases expressed as percentages whereas these are absolute numbers. why the difference? choosing one or the other (percentages, perhaps) would make it easier for readers to put the performance on the different type into a uniform context.

The reviewer is right. Text has modified in the revised version, according to the reviewer's suggestion:

“The agreement is remarkable for the PollutedSmoke cases, with occurrence ratio of 43% and 47% for NATALI and EARLINET Mahalanobis distance-based typing algorithm, respectively, whilst the CleanContinental cases showed a difference with 18% and 24% for NATALI and EARLINET Mahalanobis distance-based typing algorithm, respectively.”

47) Line 10 page 9 why is this not also true for NATALI? (I'd think the definition of the aerosol classes would be "of paramount importance" for every aerosol classification scheme)

The reviewer is right. The text has been modified to:

“Additional to the ability of the two automated algorithms to attribute the correct classification to the aerosol layers, also the uncertainties of the input measurements have to be taken into account. Both algorithms are unable to return aerosol type for not perfect calibrated input data, i.e., when one or more intensive optical properties are not perfect.”

48) Line 14 page 9 and perhaps because the dust classes are defined similarly for the two different schemes???

This is not true. As we can see from Table 3, the mean aerosol optical properties of the reference aerosol types used on the two automated algorithms, are quite different for the dust category. Higher values of Lidar ratio at 532nm are allowed for EARLINET Mahalanobis distance typing algorithm in identifying dust particles (48-62 sr), considering all dust-like aerosol types as one category. Therefore, a number of Dust recognition are attributed to PollutedSmoke particles (6 cases) or CleanContinental particles (1 case) in the NATALI output.

49) Line 22 page 9 this is a nice summary that should be introduced much earlier in the discussion

The reviewer is right. The table is provided earlier in the manuscript (In Section 3.3) and the following sentence is added in the methodology part:

“Table 2 lists the classified aerosol types of the above study and Table 3 lists the mean aerosol optical properties of the reference aerosol types defined by the two algorithms.”

50) Line 24 page 9 please explain how the model parameters are converted into synthetic lidar measurements. e.g., were realistic amounts of random noise added? how about systematic errors, like lidar calibration? were Ångström exponents used to convert 350 nm and 550 nm model parameters into 355 and 532 lidar measurements?

The ability of the ANNs to retrieve the aerosol type depends strongly on the physical content and the uncertainty of the optical inputs as well as on the structure of the ANN and the training process, including the extent of the data set used for this purpose. To create a consistent picture of the aerosol types, an aerosol model representing the optical properties of different aerosol was developed. The relative errors considered here are 50 % for the extinction, 20 % for the backscattering and 30 % for the depolarization. Note that these values were assumed to be inclusive to mimic high-precision but also moderate-precision retrieved parameters. The synthetic database was developed using the aerosol model built for 350, 550, and 1000 nm sounding wavelengths, based on the 61 wavelengths of OPAC (Optical Properties of Aerosols and Clouds) software package (Hess et al., 1998), for which the microphysical characteristics of the aerosols are available from GADS (Global Aerosol Data Set). These wavelengths are then re-scaled to the usual lidar wavelengths (i.e., 355, 532 and 1064 nm), that are used as input to NATALI, using an average Ångström exponent equal to one.

51) Line 25 page 9 that the EARLINET Mahalanobis algorithm failed to make a typing decision 39% of the time tells me that the underlying models are very different. so different, in fact, that I question the value of comparing the outputs of the two different classification schemes.

See our response to comment 45.

52) Line 28 page 9 the authors need to explain how the table values are computed; that is, what quantities are used to compute the percentages shown? since each column adds up to 100%, and since the input data comes from the NATALI models (and hence the "true" type for each sample is known), should we assume that the denominator in each column is the number of NATALI samples for a given type, and the numerator is the number of Papagiannopoulos classifications? I'd much prefer to read the authors' unambiguous explanation, rather than relying on my own assumptions about how the values were derived.

The reviewer is right and the text has been modified to:

"Each percentage is defined as follows. The denominator in each column is the number of NATALI samples for a given type, and the numerator is the number of EMD classifications for the same sample."

53) Line 3 page 10 since NATALI failed to classify 36 of 116 cases, this "very good agreement" applies only to ~2/3 of the full data set. the overall agreement between the two methods is best characterized as poor.

NATALI classified 82 of the total 110 layers (75% of the total number of cases) and 28 were flagged as Unknown layers. These missed cases are based mostly on retrievals with larger uncertainties of the optical parameters and the calculated uncertainties, that not allowed ANNs to return any aerosol type.

Indeed, the agreement can be characterized as a poor one, as it showed here, but we should keep in mind that with the used criteria in the NATALI software, the algorithm returns only the retrievals that pass a number of quality checks, so as to make sure (as possible) that the output is trustworthy. See our detailed response to comment 45.

The revised text is modified accordingly.

54) Line 4 page 10 again, it's essential that the authors explain how the percentages shown in this figure are calculated. in particular, please explain how the percentages in figure 5 relate to the sample sizes shown in figure 3.

The reviewer is right. Figure percentages are replaced by numbers, so to be consistent with the sample sizes of Figure 3. The information now is more straightforward and a paragraph is added explaining the numbers.

55) Figure 1 since the Papagiannopoulos algorithm uses AE355/1064, I would expect that quantity to be plotted here. it's a bit misleading to show 14 possibilities here when the NATALI algorithm in this study is constrained to identifying only 4 types.

please add text explaining why the probabilities do not add up to one.

The reviewer is right. The BAE 355/1064 is provided in the plot of the revised version. Also, the labels of the NATALI output were changed to the low resolution output. The text explaining the probabilities also added and given below:

“The plotted probabilities are assigned to each calculated distance and are different from the normalized ones that used as a screening criteria (and add up to 100).” See also comment on 28.

56) it's a bit misleading to show 14 possibilities here when the NATALI algorithm in this study is constrained to identifying only 4 types

The reviewer is right. The labeled categories are corrected in the plot. But, as explained above, the NATALI low resolution typing, consists of 5 types (+ the Unknown category and the N/A). See also, Table 2.

57) Figure 4 since each column adds to 100%, it would be very helpful (essential, even?) to know the number of samples in each column. are the samples uniformly distributed among the NATALI aerosol types

The number of samples in each column are added to the text.

“Fair agreement is found for the NATALI synthetic datasets of Dust and PollutedSmoke categories. In particular, for Dust reached agreement of 83.7% (100166 typed over 119650). This result is very good considering that up to 30% of the aerosol class contains mixtures. For the PollutedSmoke, the majority of the synthetic data, corresponding to 48180 instances from a total typed of 82041 (58.7%) is classified as PollutedSmoke. For the CleanContinental type, the agreement is poor (15%), as only 7605 over 52058 were correctly classified and the majority of the instances (36552) is typed as Dust (70%). This Dust-CleanContinental mismatch is confirmed also by the 11.8% of the NATALI reference dust cases that are typed as Clean Continental by the EARLINET Mahalanobis distance-based typing method. Furthermore, 7901 of the instances (15%) are classified as PollutedSmoke. Finally, the agreement for Maritime is 23% of the cases (1319 from a total of 5658 cases), indicating the different class definitions as well as the big spread of the NATALI maritime model that causes the EARLINET Mahalanobis distance-based typing algorithm to misclassify those data.”

There are some typo that are corrected accordingly:

Line 15 page 4: ‘at 355 nm’ is added

Line 20 page 4: ‘discriminate’ replace ‘detect’

Line 2 page 6: ‘available’ is added

Line 3 page 11: 'reasonably close' replace 'high enough'

Line 34 page 11: 'Furthermore' is corrected

The following references have been added in the revised manuscript

Amiridis, V., Marinou, E., Tsekeri, A., Wandinger, U., Schwarz, A., Giannakaki, E., Mamouri, R., Kokkalis, P., Biniotoglou, I., Solomos, S., Herekakis, T., Kazadzis, S., Gerasopoulos, E., Proestakis, E., Kottas, M., Balis, D., Papayannis, A., Kontoes, C., Kourtidis, K., Papagiannopoulos, N., Mona, L., Pappalardo, G., Le Rille, O., and Ansmann, A.: LIVAS: a 3-D multi-wavelength aerosol/cloud database based on CALIPSO and EARLINET, *Atmospheric Chemistry and Physics*, 15, 7127–7153, doi:10.5194/acp-15-7127-2015, <http://www.atmos-chem-phys.net/15/7127/2015/>, 2015.

D'Amico, G., Amodeo, A., Baars, H., Biniotoglou, 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, <https://doi.org/10.5194/amt-8-4891-2015>, 2015.

Giannakaki, E., Pfüller, A., Korhonen, K., Mielonen, T., Laakso, L., Vakkari, V., Baars, H., Engelmann, R., Beukes, J. P., Van Zyl, P. G., Josipovic, M., Tiitta, P., Chiloane, K., Piketh, S., Lihavainen, H., Lehtinen, K. E. J., and Komppula, M.: One year of Raman lidar observations of free-tropospheric aerosol layers over South Africa, *Atmospheric Chemistry and Physics*, 15, 5429–5442, doi:10.5194/acp-15-5429-2015, 2015.

Granados-Muñoz, M.J., Navas-Guzmán, F., Luis Guerrero-Rascado, J., Antonio Bravo-Aranda, J., Biniotoglou, I., Nepomuceno Pereira, S., 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., José Olmo, F., Papayannis, A., Pappalardo, G., Rodríguez, A., Schepanski, K., Sicard, M., Vukovic, A., Wandinger, U., Dulac, F., 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 (11), pp. 7043-7066., 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, <http://www.atmos-chem-phys.net/13/2487/2013/>, 2013.

Haarig, M., Ansmann, A., Baars, H., Jimenez, C., Veselovskii, I., Engelmann, R., and Althausen, D.: Depolarization and lidar ratios at 355, 532, and 1064 nm and microphysical properties of aged tropospheric and stratospheric Canadian wildfire smoke, *Atmos. Chem. Phys.*, 18, 11847-11861, <https://doi.org/10.5194/acp-18-11847-2018>, 2018.

Hess, M., Koepke, P., and Schult, I.: Optical properties of aerosols and clouds: The software package OPAC, *B. Am. Meteorol. Soc.*, 79, 831–844, [https://doi.org/10.1175/1520-0477\(1998\)0792.0.CO;2](https://doi.org/10.1175/1520-0477(1998)0792.0.CO;2), 1998.

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, <https://doi.org/10.5194/acp-12-2229-2012>, 2012.

Mona L., Kahn R., Papagiannopoulos N., Pappalardo G., Holzer-Popp T.: Aerosol typing—key information from aerosol satellite measurements, European Aerosol Conference, Milan (Italy) September 6-11, https://geko.prometing.it/abstract/DEF/Poster/2AAP_P094.pdf, 2015.

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, PhD 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, *Atmos. Chem. Phys.*, 12, 3115–3130, <https://doi.org/10.5194/acp-12-3115-2012>, 2012.

Shin, S.-K., Tesche, M., Kim, K., Kezoudi, M., Tatarov, B., Müller, D., and Noh, Y.: On the spectral depolarisation and lidar ratio of mineral dust provided in the AERONET version 3 inversion product, *Atmos. Chem. Phys.*, 18, 12735-12746, <https://doi.org/10.5194/acp-18-12735-2018>, 2018

Yorks J.E., Palm S.P., McGill, M.J., Hlavka D.L., Hart, W.D., Selmer P.A., and Nowottnick E.: CATS Algorithm Theoretical Basis Document, Level 1 and Level 2 Data Products, https://cats.gsfc.nasa.gov/media/docs/CATS_ATBD.pdf, 2016.

Interactive comment on Comparison of two automated aerosol typing methods and their application on an EARLINET station” by Kalliopi Artemis Voudouri et al.

Anonymous Referee #2

We would like to thank the reviewer for his/her fruitful comments that led to the improvement of the manuscript. In the following, answers to comments are reported just below each related comment. When needed, the part of the manuscript we modified or added to the old version, is reported. In the revised version of the paper all comments have been extensively taken into account discussing more aspects of the comparison between the two typing techniques (motivation, agreement and disagreement of the methods). Some text has been added about this and it is reported below into the replies to specific comments.

1) Title ‘Thessaloniki’

We keep the more generic title, since the concept of the paper is not to discuss explicitly the aerosol types observed in Thessaloniki, but is more about discussing an approach for evaluating the typing methods.

2) Line 13-14 ten : marine, dust, polluted continental / smoke, clean continental, polluted dust, elevated smoke, dusty marine, PSC aerosol, volcanic ash, sulfate/other

We thank the reviewer for his/her correction. The text has been modified to:

“The CALIPSO mission uses a decision-tree based on lidar profiles and external data (Omar et al., 2009) in order to classify the aerosol load in ten aerosol subtypes, i.e., marine, dust, polluted continental / smoke, clean continental, polluted dust, elevated smoke, dusty marine, PSC aerosol, volcanic ash, sulfate/other.”

2) Line 6 page 3 You can also provide literature for other European countries (eg Greece)

The text has been modified to: “Simultaneous observations of desert dust and ash particles were made during the Eyjafjallajokull volcanic eruption in 2010 and the methodology for the type discrimination was presented by Papayannis et al. (2012); Mona et al. (2012); and Pappalardo et al. (2013).”

3) Line 10 page 3 in the abstract you are refer to: dust, maritime, polluted smoke and clean continental. Please be consistent.

The reviewer is right. The text has been modified to:

“The two automatic aerosol classification methods and the methodology used to characterize the layers in four basic aerosol types (i.e., Dust, PollutedSmoke, Maritime, CleanContinental) are presented in Sect. 3”.

4) Line 24 page 3 Please provide a more recent paper that actually describe the recent status of THELISYS

The text has been modified to:

“A detailed description of THELISYS can be found in Siomos et al. (2018a) and Siomos (2018c).”

5) Line 2 page 4 Here you should explain the two different options of the aerosol typing algorithms (with and without δ)

Both automated algorithms can ingest depolarization information, however in our study we did not consider this possibility. The reason for this is the depolarization ratio unavailability. The next sentence text is changed to:

“The input module requires optical properties profiles as those measured by EARLINET stations, namely the aerosol extinction coefficient and the aerosol backscatter coefficient profile. Optionally, the linear particle depolarization profile at 532 nm can be provided, so as to allow a better classification and to increase the number of classified aerosols types.”

A complete description of the two different options of the aerosol typing algorithms (with and without δ), is given in Section 3.1 and Section 3.2. See below the paragraphs modified in the revised version:

“Depending on the availability of the particle linear depolarization ratio and the quality of the provided lidar profiles, the derived typing can be either of high resolution (AH), or low resolution with depolarization (AL) or low resolution without depolarization (BL). Pure aerosols categories, and even mixtures of three aerosols types can be obtained from the NATALI algorithm. In the high resolution typing, 14 aerosol types can be distinguished (i.e., Continental, ContinentalPolluted, Dust, Maritime/CC, Smoke, Volcanic, Coastal, CoastalPolluted, ContinentalDust, ContinentalSmoke, DustPolluted, MaritimeMineral, MixedDust and MixedSmoke) when the quality of the provided optical products is high enough. In the low resolution typing, 6 predominant aerosol types can be provided but with high uncertainty (i.e, Continental, Continental polluted, Smoke, Dust, Maritime and Volcanic). The low resolution typing provides 5 predominant aerosol types (Dust, ContinentalPolluted, Smoke, Continental, Maritime) either pure or mixed, when the depolarization information is not provided.”

“The algorithm applies the Mahalanobis distance classifier (Mahalanobis, 1936) to classify observations into maximum 8 (Dust, Volcanic, Mixed Dust, Polluted Dust, Clean Continental, Mixed Marine, Polluted Continental, Smoke) and minimum 4 (Dust, Maritime, PollutedSmoke, CleanContinental) aerosol classes, considering the needs of each user and the availability of the intensive properties.”

Also, the Table 2 added in the revised version, contains all the necessary information, describing both the high and low resolution aerosol typing provided by each algorithm and the low resolution that is used in our study.

6) Line 20 page 4 What about the particle depol?

The reviewer is right, therefore the following sentence is added in the revised manuscript:

“Optionally, the linear particle depolarization profile at 532 nm can be provided, so as to allow a better classification and to increase the number of classified aerosols types.”

7) Line 25 page 4 In the previous paragraph you mention that only intensive optical properties are used, but here the backscatter is used for the identification of the layers

Yes, the reviewer is right. The intensive optical properties are used for the typing, but the backscatter at 1064 nm is used only for the identification of the aerosol layers.

8) Line 6 page 5 How is high enough defined?

We can talk about high quality of lidar products, when the values of the intensive optical parameters are between acceptable limits (see Table 1) and the relative error of each intensive optical parameter is lower than 50%.

9) Line 7 page 5 you mean low resolution with depol information

Yes, the reviewer is right. The low resolution (AL) typing corresponds to the low resolution with depolarization information.

The text is modified accordingly: In the low resolution typing (AL), 6 predominant aerosol types can be provided but with high uncertainty (i.e, Continental, ContinentalPolluted, Smoke, Dust, Maritime and Volcanic). The low resolution typing (BL) provides 5 predominant aerosol types (Dust, ContinentalPolluted, Smoke, Continental, Maritime) either pure or mixed, when the depolarization information is not provided.

10) Line 8-9 page 5 Here the types are 5, in the comparison you refer to 4 aerosol types...

The reviewer is right. The methodology section is changed accordingly:

“The NATALI typing was performed in the low resolution typing configuration (5 predominant aerosol types - Dust, Smoke, Continental Polluted, Continental and Maritime) since particle linear depolarization ratio measurements for Thessaloniki were not available for the study period. In what follows, we merged the output types from NATALI that tend to reflect the same aerosol characteristics, and hence we evaluate the corresponding effects on the prediction rate of the algorithms. Thus, the smoke and the polluted continental categories were grouped into the more generic type of small particles with high lidar ratio values. The selection of four main aerosol classes stems from the availability of intensive properties, the difficulty in deriving a confident classification without particle linear depolarization ratio and the difficulty in discriminating polluted continental and smoke particles that reveal the same type characteristics. Regardless, the aerosol classes describe the major aerosol components. The identified layer boundaries from NATALI are used as input in the EARLINET Mahalanobis distance-based typing algorithm. Considering the aforementioned typing merging, the EARLINET Mahalanobis distance-based typing algorithm was set to classify observations into 4 aerosol classes: CleanContinental, Dust, Maritime and PollutedSmoke.”

Also, the following sentence is added in the last paragraph of the methodology section:

“The idea here is to compromise: i) the resolution (low) of the automatic classification owing to the availability of the optical properties (i.e., 3+2 lidar configuration), and, ii) the type definition, which

does include the wide spectrum of the aerosol types provided by the two automated typing techniques.”

Table 2 is added and lists the classified aerosol types of the above study.

11) Line 33 page 5 Here the classes are 8

The paragraph is reworded and reads as follows:

“Finally, the assessment of the predictive performance of the algorithm was tested on a testing dataset. For this purpose, 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 testing dataset comprised of 47 layers, 21 of which yielded depolarization ratio values. The performance of the algorithm was checked for each of the grouping classes (i.e., 8,7,6,5,4) and the predictive accuracy of the algorithm increased up to 90% when the aerosol classes that tend to reflect the same optical properties values were combined into 4 (Dust, Maritime, Polluted Continental + Smoke, Clean Continental) without providing the information of the depolarization. The study concluded that the fewer aerosol classes (i.e., 4, 5, 6 classes) could provide a successful prediction accuracy, even without depolarization values, but, nonetheless, a coarser and less insightful classification.”

12) Line 2 page 6 I don't understand. In this study? Increase to 90% compared to what? In general? please provide reference

Papagiannopoulos et al. (2018), checked the performance of the algorithm for each of the grouping classes and the predictive accuracy of the algorithm increased up to 90% (See Figure 10 in Papagiannopoulos et al., 2018) when the aerosol classes that tend to reflect the same optical properties were combined into 4 without providing the information of the depolarization (compared to 80% when depolarization ratio information was included). The study, also, showed that without depolarization ratio information, the accuracy of the model increases with decreasing number of classes, providing however a coarser classification. Instead, the training of the classification with depolarization measurements enhances the predictability strength and can provide finer aerosol classification (for 8 classes).

13) Line 3-5 page 6 So, you have 90% accuracy without using the depol information?

The reviewer is right. In general, the particle linear depolarization ratio increases the ability for correctly predicting the aerosol type. However, Papagiannopoulos et al. (2018) showed that (Figure 10), without depolarization ratio information, the accuracy of the model increases with decreasing number of classes (90% for the 4 classes), providing however a coarser classification.

14) Line 3-5 page 6 I understand that you have an accuracy of 90% for 4 classes. But fewer aerosol classes (I suppose fewer than 8..).

The reviewer is right. Papagiannopoulos et al. (2018) analyzed the predictive accuracy of the algorithm when compared to manually analyzed data for the different aerosol classes in both the cases in which

the depolarization information is available, for 8, 7, 6, 5 and 4 classes. This highlights that the typing in multiple classes and the typing accuracy are two conflicting aspects. The number of classes as well as the typing accuracy depends on the specific needs. This could be an approach for the specific user to select the appropriate balance each specific application. Another possibility is to find a compromise between degrading accuracy and gaining insight into the aerosol type.

15) Line 16 page 6 I suppose that you merged Smoke with Continental polluted. Please clarify this in the text

The reviewer is right. This type combination was not clearly stated clearly in the text. The selection of four aerosol classes stems from to the difficulty to distinguish aerosol types of aerosols that reflect the same aerosol characteristics. Besides, the particle depolarization ratio is not available for this study, which constitutes a powerful aerosol type discriminator. Therefore, we merged Smoke with Continental polluted. Following this comment, we added a table (Table 2) with the aerosol classes for both the High and Low resolution (without depolarization) mode for both automatic typing techniques and the aerosol types used in this study. Additionally, Papagiannopoulos et al. (2018) showed that in the testing phase of the algorithm, that the Dust classification showed a high success rate, whilst the aerosol types that performed worse were the smoke and polluted continental aerosol. However, when these two categories were combined into a single aerosol class, the correct prediction increased.

So, the following paragraph is added in the revised version of the text:

“In what follows, we merged the output types from NATALI that tend to reflect the same aerosol characteristics, and hence we evaluate the corresponding effects on the prediction rate of the algorithms. Thus, the smoke and the polluted continental categories were grouped into the more generic type of small particles with high lidar ratio values. The selection of four main aerosol classes stems from the availability of intensive properties, the difficulty in deriving a confident classification without particle linear depolarization ratio and the difficulty in discriminating polluted continental and smoke particles that reveal the same type characteristics. Regardless, the aerosol classes describe the major aerosol components. The identified layer boundaries from NATALI are used as input in the EARLINET Mahalanobis distance-based typing algorithm. Considering the aforementioned typing merging, the EARLINET Mahalanobis distance-based typing algorithm was set to classify observations into 4 aerosol classes: CleanContinental, Dust, Maritime and PollutedSmoke.”

16) Line 21-28 page 6 Maybe a simple sketch is useful here.

We excluded this paragraph and instead section 3.1 was changed to the following one, in order to provide all the necessary information.

“The identification of the most probable aerosol type is then made through a voting procedure, using the results from the three ANNs interrogated. Over 50000 aerosol synthetic data have been used to train the ANN and identify the better ANNs to classify the aerosols type from multiwavelength lidar data. The capability of ANNs to resolve the overlapping clusters of the intensive optical parameters is used on NATALI algorithm. The answer is selected based on a statistical approach. The selected types

of ANNs classify the aerosols based on the response with high i) confidence (i.e. the probability of having one of the aerosol types) and ii) stability over the uncertainty range (i.e., the percentage of agreement for values between error limits). Therefore, answers with low confidence are filtered out and NATALI returns the 'Unknown' type. In this study, we select the confidence level for the output retrievals higher than 0.9 (minimum accepted confidence) and the minimum agreement threshold as default (i.e., 0.25), so as to make sure (as possible) that the output typing is trustworthy."

17) Line 31 page 6 you mean clean continental?

We thank the reviewer for his/her comment. The text has been modified to:

"and small particles with medium lidar ratios (i.e., clean continental particles, with mean lidar ratios, for 355 and 532 nm, 50 ± 8 and 41 ± 6 sr, respectively)."

18) Line 31 page 6 I think here is necessary to give number what is low / medium and high lidar ratios and what is large and small particles

Type specific values are added to the text.

"Consequently, the lidar classification scheme consists of the main classes: (i) large particles with medium lidar ratios (i.e., dust-like particles, with mean lidar ratios of 58 ± 7 and 48 ± 55 sr for 355 nm and 532 nm, respectively, Groß et al. (2013)), (ii) large particles with low lidar ratios (i.e., maritime particles, with mean lidar ratio at 355nm and 532nm of 18 ± 4 and 18 ± 2 sr, respectively, Groß et al. (2011)), (iii) small particles with high lidar ratios (i.e., pollution and/or smoke particles, the smoke mean lidar ratio values present values of 81 ± 16 and 78 ± 11 sr for 355 and 532 nm, respectively – and the polluted continental values succeed with 69 ± 12 and 63 ± 13 sr for 355 and 532 nm, respectively, Amiridis et al., 2009; Baars et al., 2012) and (iv) small particles with medium lidar ratios (i.e., clean continental particles, with mean lidar ratios, for 355 and 532 nm, 50 ± 8 and 41 ± 6 sr, respectively). Generally, desert dust layers have optical properties that are considerably different from the other types, thus they are easily identified. Their big size leads to low Ångström exponent values and the reported lidar ratio at 355nm ranges from 47 to 58 sr for Thessaloniki (Siomos et al., 2018). PollutedSmoke particles, are also easily identified, as they are highly absorbing particles, with high lidar ratio values. CleanContinental categorization is not completely straightforward, because the continental particles can be attributed to different subcategories (i.e., local, continental polluted or mixtures). In general, the CleanContinental cases are typically elevated layers, i.e. layers not related to the local atmospheric boundary layer where the pollution and anthropogenic contribution would mean more absorbing particles and therefore labeled as PollutedSmoke aerosol. Continental particles present low lidar ratio values, (i.e., 20–40 sr) and values of Ångström exponents around 1.0–2.5. The highest values appear during summer period in Thessaloniki (Siomos et al., 2018)."

19) Line 1 page 7 I would call this category Mixed Dust and not Dust

The reviewer is right. Dust category can be either pure dust, or mixed dust, or polluted, or volcanic. But, as mixed Dust and Polluted Dust are different definitions for NATALI typing scheme, we keep label

the category with the generic 'Dust'. This is specified in Table 2, which lists the classified aerosol types of the above study.

20) Line 2 page 7 So, is there any layer Clean Continental in the urban city of Thessaloniki?

Clean continental type is defined in the CALIPSO scheme background aerosol and as a consequence, deemed not to be influenced by urban pollution. However, these conditions are probably not realistic for the European continent. So, this category was revised for the applicability to Europe in the two typing schemes, as the clean continental aerosol over Europe is a mixture of anthropogenic pollution with particles from natural sources, but with a predominance of no-anthropogenic aerosol resulting a low lidar ratio observed value (Papagiannopoulos et al., 2018).

As reported in the text many layers were typed as Clean Continental over Thessaloniki and these are typically elevated layers (90% of the detected CleanContinental cases are found above 2km). This result is compliant with the CALIPSO scheme at global level and with results about Thessaloniki site characterization done by Siomos et. al (2018). In their paper Siomos et al. (2018b), used data from a double monochromator Brewer spectrophotometer and a sunphotometer in order to classify aerosol cases during the period 2007-2017 in Thessaloniki, in the following categories: Water soluble, Black Carbon, Dust, Sea Salt and mixed. They found that the pure Water Soluble category (which can be related to the Clean Continental consisting of mainly of water soluble particles) correspond to 29.1% of the cases, which is in fair agreement with results reported in Fig.3.

We added a sentence about this in the revised version of the paper:

“CleanContinental categorization is not completely straightforward, because the continental particles can be attributed to different subcategories (i.e., local, continental polluted or mixtures). In general, the CleanContinental cases are typically elevated layers, i.e. layers not related to the local atmospheric boundary layer where the pollution and anthropogenic contribution would mean more absorbing particles and therefore labeled as PollutedSmoke aerosol. Continental particles present low lidar ratio values, (i.e., 20–40 sr) and values of Ångström exponents around 1.0–2.5. The highest values appear during summer period in Thessaloniki (Siomos et al., 2018).”

21) Line 3 page 7 I suppose that this category is pure maritime+mixed maritime

True. For making clearer this point, we added a table (Table 2) describing each aerosol category.

22) Line 7 page 7 even without depolarization ratio?and even without back-trajectory analysis?

Usually dust can be identified, since the optical properties of dust particles are quite different from the other three classes. Even though we do not consider depolarization ratio information, the size information (i.e., Angstrom exponent) and physical and chemical properties (i.e., lidar ratio) are usually sufficient to discriminate Dust. Nicolae et al. (2018) and Papagiannopoulos et al. (2018) have demonstrated this possibility. Backward trajectory analysis is a well-established tool for the source identification, however automated classification algorithms for lidars, nowadays, can operate stand-alone and provide robust results (e.g., Omar et al., 2009; Burton et al., 2012; Nicolae et al., 2018;

Papagiannopoulos et al., 2018). Indeed, support from back-trajectory analysis and/or model simulations are necessary for an independent evaluation of the stand-alone typing algorithms and this is the reason we demonstrate in our paper certain case studies (section 4.1, 4.2, 4.3)

23) Line 7 page 7 Please provide numbers

Groß et al. (2011) for example reported values of 0.06 ± 0.21 and Tesche et al. (2009b) reported values of 0.19 ± 0.20 . But, it should be pointed out that different values reported can be related to the different transportation paths of the load.

24) Line 10 page 7 Small? I suppose you are referring to aged smoke or polluted continental. Fresh smoke is not small

Yes, the reviewer is right. Smoke particles yield a wide range of optical properties. The transportation path, the aging, the burning material, the hygroscopic growth, and the height of injection are some of the parameters that can affect the observed optical properties. For instance, smoke layers observed by EARLINET systems over Europe in the summer of 2017 indicated high depolarization ratio (e.g., Haarig et al., 2018). The mechanisms for this are still yet to be understood. Here, we merged Smoke with Continental polluted, and the characteristics correspond to the merged class. We would like, also, to comment that we had to i) compromise the resolution of the automatic classification owing to the availability of the optical properties (i.e., 3+2 lidar configuration), and, ii) the type definition, which does include the wide spectrum of the aerosol types as stated above.

25) Line 32 page 7 Please correct the label of Natali in Figure 1, it should have run for 5 (merged to 4) classes

Following this comment, the plot is changed and the NATALI labels, now, are 5 (+ 2 categories of 'Unknown' and 'N/A' typing), corresponding to the categories of the low resolution typing.

26) Line 1 page 8 Please be consistent: use either Layer A, B or Layer 1, 2

The plot and the text is changed accordingly, using the labels of Layer 1, 2.

27) Line 2 page 8 'and Angstrom exponent'

Following this comment, the text is changed to: 'The stability of the lidar ratio and Ångström exponent values could be considered as an indicator of homogeneity and small variability of the aerosol type within the layer.'

28) Line 12 page 8 please present the fire spots in the trajectory already provided in Figure 2. Also give reference for the fires.

The fire spots are added in Figure 2. These values are in accordance with the typical biomass burning values observed over Thessaloniki. Giannakaki et al. (2010) report an annual mean lidar ratio at 355nm of 69 ± 17 sr and a mean BAE at 355-532nm of 1.7 ± 0.7 , while Siomos et al. (2018a) found the lidar ratio at 355nm ranging from 51 to 73 sr for biomass burning events.

The paragraph is reworded and reads as follows:

“These values are in accordance with the typical biomass burning values observed over Thessaloniki. Giannakaki et al. (2010) reported an annual mean lidar ratio at 355nm of 69 ± 17 sr and a mean BAE at 355-532nm of 1.7 ± 0.7 , while Siomos et al. (2018a) found lidar ratio at 355nm ranging from 51 to 73 sr for biomass burning events.”

29) Line 23 page 8 I would like to see, the application of the algorithms not only to pure dust and bb cases. To classify these pure cases is relative easy and there is not a need for aerosol classification algorithm. The authors should provide additional case studies with mixed aerosol types, or cases with complex aerosol structure and check the reliability of the algorithms.

The reviewer is right. A new plot and a paragraph (Section 4.3) is added in the revised version. The typing scheme selected and added is a more complex one and offers the opportunity to check the reliability of the algorithms in conditions where different aerosol types at different heights are observed.

30) Line 33 page 8 It is difficult to believe that the second aerosol type observed in Thessaloniki is clean continental. Have you also compared these results with satellite/model products? Clean continental is defined as 'background like' aerosols with a LR of about 30-35sr. I would speculate that the observed in this study layers would be small, non so absorbing aerosol.

See also our response to comment 20. Typically, the clean continental type presents the mixture of anthropogenic aerosols with natural sources. For instance, the manual typing made by Schwarz et. al. (2016) in the framework of EARLINET, assigns an aerosol layer as clean continental when the aerosol concentration is low by means of optical depth (the low values of lidar ratio). Based on this definition, the label “clean” can be taken literally. However, the automatic typing procedures do not take into account any extensive parameter, hence not “clean” aerosol layers might be classified as clean continental.

31) Line 3 page 9 We should talk for agreement only if the same layer is attributed to the same class. In such a figure we only new the statistical information of the occurrence of the layers, not if the classification is made right.

The reviewer is right. The paragraph is reworded and reads as follows:

“In particular, the agreement is reasonably close for the desert dust cases (10% and 17% for NATALI and EARLINET Mahalanobis distance-based typing algorithm, respectively), nevertheless, it becomes evident that the particle linear depolarization ratio could increase the ability for correctly predicting dust particles.”

32) Line 24 page 9 Please explain possible reasons for that

After the bug in the processing, the scores have changed to the following ones and the text has been modified to:

“For PollutedSmoke and CleanContinental the accuracy reached 88% and 65% respectively.”

In order to address the possible reasons of disagreement, we added the following paragraph in section 4.4 of the revised version of the manuscript and a summary table (Table 3) which provides mean aerosol optical properties of the reference aerosol types used on the two automated algorithms.

Overall, there were 15 cases (on Thessaloniki dataset) that the two methods provided different typing results. There are seven cases typed as CleanContinental aerosol by NATALI and PollutedSmoke or CleanContinental by EMD, five cases typed as PollutedSmoke by NATALI and Dust by EMD, 2 cases typed as Dust by NATALI and CleanContinental by EMD and one case type as Maritime by NATALI and CleanContinental by EMD. These mismatches are illustrated in Figure 7. In order to understand these differences we highlight some critical issues relevant the type definitions of the two methods, based on Nicolae et al., (2018) and Papagiannopoulos et al., (2018).

CleanContinental: The contribution of the Soot component in the chemical composition of the CleanContinental category, allows higher lidar ratio values at 532nm (52–53sr) in the NATALI scheme. Consequently, layers recognized by NATALI as CleanContinental ones, in the EARLINET Mahalanobis distance-based typing method are attributed as signature for PollutedSmoke or Dust (as it can be seen in the revised Figure 6 -now Figure 7).

Marine: As observations of pure maritime particles are quite scarce within EARLINET and, generally, when these particles are observed their characteristics are far from pristine, the Maritime category for the EARLINET Mahalanobis distance typing algorithm, corresponds to mixed maritime layers. This is different from the pure maritime category that NATALI identifies. Lower lidar ratio values for 532nm (19–25sr) are defined in the NATALI software for the identification for marine layers, in contrast to the higher ones (16-32sr) allowed in the EARLINET Mahalanobis distance typing algorithm.

Dust: Higher values of Lidar ratio at 532nm are allowed for EARLINET Mahalanobis distance typing algorithm in identifying dust particles (48-62sr), considering all dust-like aerosol types as one category, while NATALI allows values corresponding to more pure cases (44–49). Therefore, a number of Dust recognition are attributed to PollutedSmoke particles (6 cases) or CleanContinental particles (1 case) in the NATALI output.

PollutedSmoke: An almost perfect score was found for the PollutedSmoke category. This can be attributed to the similar reference values attributed by both typing algorithms. Higher values of Lidar ratio at 532nm are allowed for NATALI (62-92sr - ContinentalPolluted and Smoke), instead of the lower ones (52–89sr) allowed in the EARLINET Mahalanobis distance typing algorithm.”

The paragraph describing the mismatches is reworded and reads as follows:

“Although, each automated classification algorithm has important differences acknowledged above, the comparison showed an overall good agreement for the four defined aerosol classes. The convergence of the two different methods on the same type can be regarded as a signature of reliability. An almost perfect score was found for the PollutedSmoke category, following by the Dust one, given that the dust class is well defined for both typing schemes, as the physical properties of dust particles differentiate from the other three classes. Considering the mismatches in the

CleanContinental category, the agreement is good enough. By contrast, the maritime category is defined in a different way for the two automated algorithms. The EARLINET Mahalanobis distance-based typing algorithm considers maritime layers mixed with other aerosol types, whereas for the NATALI the mixing is negligible and the aerosol type refers to pure maritime aerosol. The absence of measurements for such kind of particle also did not allowed a direct assessment of the pure marine particle synthetic data into NATALI algorithm itself. The case typed as Maritime by NATALI was identified as CleanContinental: this is because of the different lidar ratio at 532nm and backscatter Ångström related values allowed in the NATALI scheme which are recognized by the EARLINET Mahalanobis distance-based typing method as signature for CleanContinental types (Table 3)."

33) Line 12 page 10 case studies

The text has been modified to:

"The prediction of the automatic classification methods in the three case studies showed consistent results when compared against manually classified EARLINET data."

34) Figure 1

S stands for Pure smoke or general for the category: polluted continental + smoke?

The reviewer is right. The PS label is added to the plot, that corresponds to the category: polluted continental + smoke.

35) Figure 2

3rd, not 3th. please correct the label, only 4 types should be appear here.

The reviewer is right. The name of the layer is corrected (3rd was replaced with 3th).

Is it pure smoke?

please correct the label, only 4 types should be appear here. Is it pure smoke?

The reviewer is right. The PS label was added to the plot, that corresponds to the category: polluted continental + smoke.

The following references have been added in the revised manuscript

References

Granados-Muñoz, M.J., Navas-Guzmán, F., Luis Guerrero-Rascado, J., Antonio Bravo-Aranda, J., Biniotoglou, I., Nepomuceno Pereira, S., 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., José Olmo, F., Papayannis, A., Pappalardo, G., Rodríguez, A., Schepanski, K., Sicard, M., Vukovic, A., Wandinger, U., Dulac, F., 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 (11), pp. 7043-7066., 2016.

Groß, S., Tesche, M., Freudenthaler, V., Toledano, C., Wiegner, M., Ansmann, A., Althausen, D., and Seefeldner, M.: Characterization of Saharan dust, Maritime 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:10.1111/j.1600-0889.2011.00556.x, 2011.

Haarig, M., Ansmann, A., Baars, H., Jimenez, C., Veselovskii, I., Engelmann, R., and Althausen, D.: Depolarization and lidar ratios at 355, 532, and 1064 nm and microphysical properties of aged tropospheric and stratospheric Canadian wildfire smoke, *Atmos. Chem. Phys.*, 18, 11847-11861, <https://doi.org/10.5194/acp-18-11847-2018>, 2018.

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, <https://doi.org/10.5194/acp-12-2229-2012>, 2012.

Papayannis A., Mamouri R-E., Vassilis A., Giannakaki E., Veselovskii I., Kokkalis P., Tsaknakis G., Balis D., Kristiansen Nina, Stohl Andreas, Korenskiy Mikhail, Allakhverdiev Kerim : 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 - ATMOS ENVIRON.* 48. 10.1016/j.atmosenv.2011.08.037, 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., D'Amico, G., Comerón, A., Mona, L., Alados- Arboledas, L., Amodeo, A., Baars, H., Baldasano, J. M., Belegante, L., Biniotoglou, 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., Haeferle, 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, *Atmos. Meas. Tech.*, 8, 4587–4613, <https://doi.org/10.5194/amt-8-4587-2015>, 2015.

Siomos N., Fountoulakis I., Natsis A., Karagiozidis D., Drosoglou T., Voudouri K.A., Bais A. and Balis D.: Aerosol classification based on the synergy of Brewer spectrophotometer and sunphotometer measurements, 14th International Conference on Meteorology, Climatology and Atmospheric Physics, <http://comecap2018.gr/wp-content/uploads/2019/03/COMECAP-BOA-final-revised.pdf>, 2018b.

Siomos, N., *Climatology and classification of the optical and microphysical properties of aerosols using ground - based remote sensing techniques (Doctoral dissertation)*. Retrieved from <http://hdl.handle.net/10442/hedi/45104>, 2018c.

Comparison of two automated aerosol typing methods and their application on an EARLINET station.

Kalliopi Artemis Voudouri¹, Nikolaos Siomos¹, Konstantinos Michailidis¹, Nikolaos Papagiannopoulos^{2,3}, Lucia Mona², Carmela Cornacchia², Doina Nicolae⁴, and Dimitris Balis¹

¹Laboratory of atmospheric physics, Physics Department, Aristotle University of Thessaloniki, Greece

²Consiglio Nazionale delle Ricerche, Istituto di Metodologie per l'Analisi Ambientale (CNR-IMAA), Tito Scalo (PZ), Italy

³CommSensLab, Dept. of Signal Theory and Communications, Universitat Politècnica de Catalunya, Barcelona, Spain

⁴National Institute of R&D for Optoelectronics (INOE2000), Magurele, Romania

Correspondence: Kalliopi Artemis Voudouri (kavoudou@auth.gr)

Abstract. In this study we ~~compare two automatic~~ apply and compare two algorithms for the automated aerosol type characterization of the aerosol layers derived from ~~a Raman lidar and we test their application over a specific environment in continental Thessaloniki~~ Raman lidar measurements over the EARLINET station of Thessaloniki, Greece. Both automated aerosol typing type characterization methods base their typing on lidar derived aerosol intensive properties. The methodologies are briefly described and ~~the application on two case studies is presented. The results are checked for their consistency with satellite products and model simulations. Further application of~~ their application on three distinct cases is demonstrated and evaluated. Then the two classification ~~tools was conducted. The evaluated dataset~~ schemes were applied in the automatic mode on a more extensive dataset. The dataset analyzed corresponds to ACTRIS/EARLINET (European Aerosol Research Lidar NETwork) Thessaloniki data acquired during the period 2012-2015. ~~80 layers were successfully~~ Seventy one layers out of 110 (percentage of 65%) were typed by both ~~algorithms and assigned to four techniques~~ and 56 of these 71 layers (percentage of 79%) were attributed to the same aerosol type. However as shown the identification rate of both typing algorithms can be changed regarding the selection of appropriate threshold criteria. Four major types of aerosols are considered in this study: Dust, Maritime, PollutedSmoke and ~~Clean Continental~~ CleanContinental. The analysis showed that the two algorithms ~~are in a very good agreement,~~ when applied to real atmospheric conditions, ~~with an agreement percentage of 88.8% for Dust, 93.7% for PollutedSmoke and 70% for CleanContinental. The Maritime category was the one with the largest spread. These differences are~~ provide typing results that are in good agreement regarding the automatic characterization of PollutedSmoke, while there are some difference between the two methods regarding the characterization of Dust and CleanContinental. These disagreements are mainly attributed to differences in defining the aerosol types for the two methods. The overall consistency the definitions of the aerosol typing types between the two ~~automatic procedures despite the different aerosol type definition, allows their applicability to~~ lidar data for characterization purposes. The jointly characterization shows the high degree of confidence in identifying Dust and PollutedSmoke, as well as the need of further investigation for Maritime and Clean Continental typemethods, regarding the intensive properties used and their range.

1 Introduction

~~The aerosol~~ Aerosol classification is a ~~critical topic~~ key parameter in understanding the impact of different aerosol sources on climate, weather systems, and air quality. Given that, aerosols are emitted from various sources, a significant number of aerosol types with different optical and microphysical properties coexist in the atmosphere. As a consequence, considerable
5 uncertainties have been raised in the quantification of aerosols in the radiative transfer calculations and their interactions with clouds, leading the interest of wide communities to improving common understanding related to aerosol types and sources.

Nowadays, aerosol typing is an interesting topic recognized at international level, as inferring the aerosol source is an important tool for identification of events, alerting of dangerous situations, definition of mitigation strategies. The International Satellite Aerosol Science Network (AEROSAT), for example, is an initiative established from different research groups around
10 the world focusing on satellite aerosol retrieval. AEROSAT identifies the need of a common background for aerosol typing and typing procedures comparison efforts as one of the main important task to be tackled into the aerosol community (Mona et al., 2015).

Thus, a number of typing schemes have been applied to classify aerosols, based on the synergistic information of the ~~scattering mechanisms~~ backscattering and extinction coefficients and the polarization state of the received light ~~of at~~ different
15 wavelengths (e.g., Ansmann and Müller, 2005, Groß 2013). For example, Hamill et al. (2016), developed an aerosol classification scheme using the optical properties derived from the AERONET (Aerosol Robotic Network) ground-based network of sun photometers. Five types of aerosols were defined based on reference clusters, characteristic of a particular type of aerosol (i.e., Maritime, Urban Industrial, Dust, Mixed and Biomass Burning). Schmeisser et al. (2017) also, used aerosol optical properties (i.e., scattering Ångström exponent, absorption Ångström exponent and single scattering albedo) derived from in situ
20 measurements in order to classify aerosols. Their typing scheme was based on three different aerosol classification methods.

Active remote sensing instruments such as ground or space-based lidars, ~~consist are~~ a key technique for characterizing aerosols as ~~it provides they can provide~~ vertically-resolved information of extensive (e.g., aerosol backscatter coefficient, aerosol extinction coefficient and volume depolarization ratio) and intensive properties (e.g., Ångström exponent, lidar ratio and particle depolarization ratio) of different aerosol types. The
25 intensive properties extensive properties depend on the aerosol concentration, whilst intensive ones are type sensitive and provide separate classification for each detected layer (e.g., Müller et al., 2007; Ansmann et al., 2011; Tesche et al., 2011; Burton et al., 2012; Pappalardo et al., 2013; Groß et al., 2013; Amiridis et al., 2015; Giannakaki et al., 2015; and Baars et al., 2016). Nevertheless, even intensive properties might not be sufficient to guarantee accurate typing, as some aerosol types (e.g., biomass burning and industrial pollution) have very similar intensive properties but are attributed to different sources and generating mechanisms. Indeed, volcanic and desert dust particles are
30 both characterized by the asphericity of the particles, observable and measured through the linear depolarization ratio (while typically the other aerosol types have small values of this parameter) and also, their intensive optical properties could be similar in ranges e.g., Ångström exponents and lidar ratio (see for example Fig. 5 in Nicolae et al., 2018; Fig. 7 in Papagiannopoulos et al., 2018 and the variety of volcanic values in Mona et al., 2012).

Aerosol typing schemes have been developed for high-resolution lidar measurements of space-borne lidars (e.g., CALIPSO, Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation, Omar et al. 2009; EarthCARE, Illingworth et al. 2015; CATS, Yorks et al., 2016), airborne - High Spectral Resolution Lidar (~~HSRL~~-measurements-~~HSRL-~~measurements (e.g., Burton et al., 2012; Groß et al., 2013) and multiwavelength Raman measurements (e.g., Nicolae et al., 2018; Papagiannopoulos et al., 2018). The CALIPSO mission uses a decision-tree based on lidar profiles and external data (Omar et al., 2009) in order to classify the aerosol load in ~~seven-ten~~ aerosol subtypes, i.e., ~~clean-continental, clean, maritime~~marine, dust, polluted continental /~~smoke, clean continental~~, polluted dust, ~~smoke, and dusty-maritime~~elevated smoke, dusty marine, PSC aerosol, volcanic ash, sulfate/other. The upcoming mission of EarthCARE will depend on the high-spectral-resolution lidar products for aerosol classification. The aerosol intensive properties at 355nm combined with observations of the ~~spectral-355nm~~ aerosol optical depth will constitute the basic input for the aerosol-type determination (Wandinger et al., 2016). An aerosol classification scheme from HSRL aerosols measurements has been introduced by Burton et al. (2012). Their technique applies an objective multivariate analysis using lidar intensive properties (i.e., the particle linear depolarization ratio at 532 nm, the particle lidar ratio at 532 nm, the backscatter-related ~~532-to-1064nm~~-532 -to- 1064nm color ratio, and the ratio of particle linear depolarization ratios at 1064 and 532 nm) and is able to ~~detect~~-discriminate 8 aerosol types: smoke, fresh smoke, urban, polluted maritime, maritime, dusty mix, pure dust and ice.

Several other approaches have been developed based on ~~of~~ the information provided by multispectral data of ground based measurements. EARLINET, for example, is a well structured network of advanced laser remote sensing stations with the main purpose of understanding the horizontal and vertical distribution of aerosols on the European scale. The EARLINET data have been extensively used not only for climatological studies (e.g., Matthias et al., 2004; Mattis et al., 2004; ~~Amiridis2005~~Amiridis et al., 2005), but also for studies on Saharan dust outbreaks (e.g., Ansmann et al., 2003; Balis et al., 2004; Papayannis et al., 2008; Tesche et al., 2009b; Mona et al., 2014; ~~Biniotoglou et al., 2015~~), volcanic eruptions (e.g., Pappalardo et al., 2004; Sawamura et al., 2012; Pappalardo et al., 2013) and biomass burning events (e.g., Balis et al., 2003; Tesche et al., 2011; Nicolae et al., 2013).

Efforts for detailed knowledge of the aerosol sources have been conducted in the framework of EARLINET, focusing on the aerosol characterization and typing. For example, Müller et al. (2007) presented for the first time a statistical analysis of lidar ratios for almost all climatically relevant aerosol types based solely on Raman lidar measurements. The analysis covered the most important aerosol types such as maritime particles, desert dust particles and aged biomass-burning smoke. Lidar depolarization measurements have also been used to differentiate between different types of aerosols, since they constitute an indicator of the sphericity of particles. Tesche et al. (2009) separate the optical properties of desert dust and biomass burning particles using multiwavelength Raman and depolarization measurements. Methods for fine and coarse mode separation for dust outbreak cases, using the 532 nm particle depolarization ratio have also been developed (e.g., ~~Ansmann et al., 2012~~Mamouri and Ansmann, 2014). Additionally, combined measurements of the lidar ratio from ground-based Raman lidars along with aerosol depolarization values and the size-sensitive Ångström exponent, were proved to be a useful tool for the separation of aerosol types as shown by Groß et al. (2011). Furthermore, the combined use of lidar observations and transport model simulations permit the discrimination of desert dust and volcanic ash particles that typically, have the same optical characteristics. Simul-

taneous observations of desert dust and ash particles were made during the ~~Eyjafjallajökull~~ Eyjafjallajökull volcanic eruption in 2010 and the methodology for the type discrimination was presented by [Papayannis et al. \(2012\)](#); [Mona et al. \(2012\)](#); and [Pappalardo et al. \(2013\)](#).

~~In this article~~ Automated aerosol typing schemes have been discussed to be implemented in the next version of the EARLINET Single Calculus Chain tool (SCC; [D'Amico et al., 2015](#)). SCC was developed for the analysis of the data of different lidar systems in an automated, unsupervised way and currently SCCv5.0.11 delivers profiles of optical aerosol properties. Future work is dedicated to the implementation of new features like profiles of intensive optical properties, detection of aerosol layer geometrical properties and calculation of the intensive optical properties per detected layer, which will further allow the classification of the observed layers into aerosol types. Therefore, the need for defining a common background on aerosol types (even nomenclature is not harmonized nowadays) and for comparing aerosol types (which is not a numerical value) derived by different methods is on the top of interest and the following questions have been raised:

- [Is there a common understanding for the aerosol types between different typing schemes?](#)
- [Can automated methods provide operationally accurate typing in automatic mode when applied in data stored in a climatological database?](#)
- 15 • [How can we compare different typing methodologies? Which parameters affect the agreement or disagreement between them and the ability to provide typing results.](#)

[Given the aforementioned status](#), a first attempt of comparing and evaluating two classification tools [developed within EARLINET](#), that provide near-real-time aerosol typing information for the lidar profiles of Thessaloniki, ~~developed within EARLINET~~ is presented. [Our aim is: \(i\) to check the performance of both supervised learning techniques in their low resolution mode, which is the case for the majority of the available measurements within EARLINET, when applied to lidar data from a station where, typically, variable mixtures of aerosols are present and \(ii\) to investigate the reasons of typing agreement and disagreement with respect to the uncertainties and the threshold criteria applied. This paper can be considered as a step towards the general objective of finding translating rules and way for quantifying differences between typing procedures.](#)

The article is structured as follows: [the](#) Thessaloniki EARLINET lidar station is presented in Sect. 2. The two automatic aerosol classification methods and the methodology used to characterize the layers in four basic aerosol types (i.e., ~~desert, smoke, maritime, continental~~ Dust, PollutedSmoke, Maritime, CleanContinental) are presented in Sect. 3. In Sect. 4 the accuracy of the algorithms is tested by comparison with pre-classified case studies ~~and an~~, [the](#) evaluation for the whole period under study is ~~also presented~~. [The presented and the results of the study are discussed. Finally, Sect. 5 contains the summary and the conclusions of this article are presented in Sect. 5.](#)

30 2 Lidar system and measurement site

The Thessaloniki lidar system (THELISYS) is operated and maintained by the Laboratory of Atmospheric Physics that is located in the Physics Department of the Aristotle University of Thessaloniki (40.5° N, 22.9° E, 50m). THELISYS is used for

the detection of aerosol particles as a part of EARLINET (Bosenberg et al., 2003; Pappalardo et al., 2014) since 2000 and now is part of the Aerosols, Clouds, and Trace gases Research InfraStructure (ACTRIS; www.actris.eu/). Systematic measurements are performed, following EARLINET's schedule (i.e., every Monday morning, and every Monday and Thursday evening after the sunset). Additional measurements are performed during Saharan dust outbreaks, smoke advection from biomass burning, volcanoes [eruption](#), and CALIPSO correlative measurements. The current setup of THELISYS includes three elastic backscatter channels at 355 nm, 532 nm and 1064 nm, two nitrogen Raman channels at 387 nm and 607 nm and two depolarization channels. These two channels have been added to measure the cross and parallel polarized signal at 532 nm, but due to technical issues the particle depolarization ratio is ~~currently not available~~ [available after August 2018](#). A detailed description of THELISYS can be found in [Amiridis-Siomos et al. \(20052018a\) and Siomos \(2018c\)](#). Data from THELISYS are regularly analysed and quality assured and are publically available at www.earlinet.org.

Thessaloniki is in a location where many different types of aerosols coexist (Amiridis et al., 2009; Giannakaki et al., 2010; Siomos et al., [20182018a](#)). Dust events are dominant during summer above 1.5 km and in autumn below 1.5 km as shown by Siomos et al. ~~,(2018(2018a)~~. Marinou et al. (2017), also, used CALIPSO data and confirmed the existence of dust plumes during advection episodes over 2 km in summer. Similarly, the most intense biomass burning episodes tend to occur during summer in the free troposphere and are probably associated to wildfires rather than agricultural fires that tend to be predominant during spring and autumn (Siomos et al., 2018). Continental layers observed over Thessaloniki station are attributed to mixtures of anthropogenic pollution and particles from natural sources and even mixtures of maritime aerosol. Therefore, Thessaloniki is well suited for aerosol typing studies and for the investigation of the performances of different aerosol typing algorithms.

3 Aerosol Typing Methods

The two automatic aerosol typing methods require only lidar data with $3\beta + 2\alpha$ ~~(+1 δ) configuration~~ [configuration \(3 backscatter and 2 extinction coefficient profiles\)](#) without any use of ancillary external information. Specifically, the typing methods make use of the [available](#) aerosol type-sensitive intensive properties. ~~Multiwavelength~~ [As multiwavelength](#) Raman lidars have the ability to measure directly aerosol extinction and aerosol backscatter coefficient profiles in several wavelengths. ~~Thus~~, a number of intensive properties can be ~~obtained~~ [These aquired with high accuracy. A number of the obtained](#) quantities do not depend on the aerosol load but they can be linked to the size, the chemical composition, and/or the asphericity of the particles. The investigation of these quantities is important to infer the aerosol type as discussed in many papers (e.g. Burton et al., 2012; Groß et al., 2013; Wandinger et al, 2016). [In general, the ability for a succesfull classification of the different aerosol types and mixtures depends on the measurments uncertainties. As large measurement uncertainties prevent a correct aerosol type separation, high quality measurements are mandatory. The EARLINET lidar data follow the quality standards that have been established, in order to make the lidar products of the different systems comparable and to be able to provide quality-assured datasets of network products \(Freudenthaler et al., 2018\). So, the lidar dataset used in this study can be considered for typing characterization.](#)

The intensive properties relevant to this study are: the extinction-related Ångström exponent (AE), the backscatter-related Ångström exponent (BAE), the ratio of the backscatter coefficients profiles (color ratios - CI), the lidar ratio (LR), and the ratio of the lidar ratios - (RLR). The respective formulas are provided in the following equations, where λ is the wavelength, z is the height, a is the aerosol extinction coefficient, and b is the aerosol backscatter coefficient.

$$5 \quad AE_{\lambda_1/\lambda_2}(z) = -\frac{\ln\left(\frac{a(\lambda_1, z)}{a(\lambda_2, z)}\right)}{\ln\left(\frac{\lambda_1}{\lambda_2}\right)} \quad (1)$$

$$BAE_{\lambda_1/\lambda_2}(z) = -\frac{\ln\left(\frac{b(\lambda_1, z)}{b(\lambda_2, z)}\right)}{\ln\left(\frac{\lambda_1}{\lambda_2}\right)} \quad (2)$$

$$CI_{\lambda_1/\lambda_2}(z) = \frac{b(\lambda_1, z)}{b(\lambda_2, z)} \quad (3)$$

$$LR(\lambda_1, z) = \frac{a(\lambda_1, z)}{b(\lambda_1, z)} \quad (4)$$

$$RLR_{\lambda_2/\lambda_1}(z) = \frac{LR(\lambda_2, z)}{LR(\lambda_1, z)} \quad (5)$$

10

The aerosol backscatter coefficients at the two wavelengths (λ_1 and λ_2) are combined to give a backscatter-related aerosol Ångström exponent - AE(λ_1, λ_2) - BAE(λ_1/λ_2). This quantity provides information about the aerosol size, and in contrast to the extinction coefficient, is associated with smaller uncertainty. This is attributed mainly to the derivative in the inversion algorithm (e.g., Ansmann et al., 2002; Freudenthaler et al., 2009), which produce larger uncertainties, while the backscatter coefficient calculated from the combination of Raman-elastic channels is less sensitive. The ratio of the aerosol extinction to backscatter coefficient is called lidar ratio - LR λ_1 - and changes largely for aerosols with different chemical and physical properties. This quantity has been valuable for aerosol characterization as demonstrated in previous studies (Müller et al., 2005; Mattis et al., 2004). The ratio of the lidar ratios - LR λ_1 /LR λ_2 - can be used to assess the spectral dependency of the different aerosol types.

It is worthwhile to mention that the particle linear depolarization ratio is an intensive property that effectively discriminates spherical and non-spherical particles in the atmosphere. However Nevertheless, the particle depolarization can be very sensitive to the calibration procedure, and the values of the calculated particle depolarization introduce also, a significant uncertainty. Given that, this quantity is not used in the aerosol typing presented here, this study assesses aerosol typing capabilities for 3+2 lidars that do not make depolarization measurements.

25

3.1 Neural network Aerosol Typing Algorithm based on Lidar data - NATALI

The NATALI (Neural network Aerosol Typing Algorithm based on Lidar data) software relies on Artificial Neural Networks (Nicolae et al., 2018). The development of this tool started in the framework of EARLINET, with the main purpose of identifying the most probable aerosol type using a combination of mean-layer intensive optical parameters (i.e., lidar ratios, Ångström exponent, color ratios) from the provided aerosol backscatter and extinction coefficient profiles of lidar systems, without any additional information. The NATALI software consists of three independent, but inter-connected modules: the input, the typing and the output module. The input module requires optical properties profiles as those measured by EARLINET stations, ~~i.e., namely the~~ aerosol extinction coefficient ~~, and the~~ aerosol backscatter coefficient ~~and particle linear depolarization ratio profiles~~ profile. Optionally, the linear particle depolarization profile at 532 nm can be provided, so as to allow a better classification and to increase the number of classified aerosols types.

In a first step, the typing module identifies the geometrical boundaries of the layers by applying the gradient method on the 1064 nm backscatter coefficient profile (Belegante et al., 2014). For every detected layer from the input module, calculations of the mean layer values of the intensive optical parameters ~~, with respect to the signal to noise ratio~~ and the associated uncertainties are performed (Nicolae et al., 2016). The extinction-related Ångström Exponent at 355nm and 532nm ($AE_{355/532}$), the backscatter-related Ångström Exponent at 355nm and 532nm ($BAE_{355/532}$), the backscatter-related Ångström Exponent at 532nm and 1064nm ($BAE_{532/1064}$), the lidar ratio at 355 nm (LR_{355}), the lidar ratio at 532 (LR_{532}), and the ratio of the backscatter coefficients profiles (color ratios) $CR_{355/532}$ and $CR_{532/1064}$, are calculated. The acceptable ranges for the calculated intensive optical parameters should be between acceptable limits (i.e., extinction-related Ångström Exponent values, backscatter-related Ångström Exponent values and color ratio values between -2 and 6 and lidar ratio values between 5 and 200), so as the quality of the input data to be ensured.

In a second step ~~, the calculations of the optical properties~~ three Artificial Neural Networks (ANN) are interrogated simultaneously regarding the aerosol type using the aforementioned optical properties calculated for each layer are used in the typing module and compared to values obtained. The ANNs have been trained using synthetic data from a specially designed aerosol model ~~with simulations from over 50000 synthetic cases of aerosols~~. Specifically, the synthetic database was developed using the aerosol model built for 350, 550, and 1000 nm sounding wavelengths, based on the 61 wavelengths of OPAC (Optical Properties of Aerosols and Clouds) software package (Hess et al., 1998), for which the microphysical characteristics of the aerosols are available from the Global Aerosol Data Set (GADS). These wavelengths are then re-scaled to the usual lidar wavelengths (i.e., 355, 532 and 1064 nm) using an average Ångström exponent equal to one. This was considered a valid assumption for all aerosol types, taking into account the small difference between the lidar and the model wavelengths. A comprehensive description of the developed aerosol model can be found in Nicolae et al., (2018).

The identification of the most ~~likely~~ probable aerosol type is then made through a voting procedure, using the results from the three ANNs interrogated. Over 50000 aerosol synthetic data have been used to train the ANN and identify the better ANNs to classify the aerosols type from multiwavelength lidar data. The capability of ANNs to resolve the overlapping clusters of the intensive optical parameters is used on NATALI algorithm. The answer is selected based on a statistical approach ~~and~~

~~considering answers with a) high confidence, and b) stable.~~ The selected types of ANNs classify the aerosols based on the response with high i) confidence (i.e. the probability of having one of the aerosol types) and ii) stability over the uncertainty range (Nicolae et al., 2018). The Ångström Exponent at 550nm and 350nm ($AE_{550/350}$), the lidar ratio at 350 nm (LR_{350}), the lidar ratio at 550 (LR_{550}), and color ratios $CR_{550/350}$ and $CR_{1000/550}$, are used for the aerosol classification i.e., the percentage of agreement for values between error limits). Therefore, answers with low confidence are filtered out and NATALI returns the 'Unknown' type. These threshold criteria are selected by the user and are expressed as percentages. Higher values of these criteria, enhance the confidence of the classification and affect the identification rate of the typed cases. In our study, we test their application and we define the confidence level for the output retrievals higher than 90% (minimum accepted confidence) and the minimum agreement threshold as default (i.e., 25%), in order to check the performance of NATALI in the highest confidence level.

Depending on the availability of the particle linear depolarization ratio and the quality of the provided lidar profiles, the derived typing can be either of high resolution (AH), or low resolution with depolarization (AL) or low resolution without depolarization (BL). Pure aerosols categories, and even mixtures of three aerosols types can be obtained from the NATALI algorithm. In the high resolution typing, 14 aerosol types can be distinguished (i.e., Continental, ContinentalPolluted, Dust, Maritime/CC, Smoke, Volcanic, Coastal, CoastalPolluted, ContinentalDust, ContinentalSmoke, DustPolluted, MaritimeMineral, MixedDust and MixedSmoke) when the quality of the provided optical products is high enough. In the low resolution typing (AL), 6 predominant aerosol types can be provided but with high uncertainty (i.e. Continental, ~~Continental polluted~~ContinentalPolluted, Smoke, Dust, Maritime and Volcanic). The low resolution typing (BL) provides 5 predominant aerosol types (i.e., Dust, ContinentalPolluted, Smoke, Continental, Maritime) either pure or mixed, when the depolarization information is not provided. Finally, the output module provides the intensive optical parameters within each layer as long as along with their mean value and the corresponding uncertainty. The complete and detailed typing procedure derived by NATALI can be found at Nicolae et al., (2016) and at Nicolae et al., (2018).

Application of NATALI on EARLINET data samples was also conducted. Observational data from the EARLINET-CALIPSO database and measurements derived by a Raman depolarization Lidar from the Romanian National Institute for Research and Development in Optoelectronics (Belegante et al., 2011) were used. The comparison between the NATALI aerosol typing module and these observational data showed ~~positive-consistent~~ results (Nicolae et al., 2018). This comparison on EARLINET data samples showed the capability of NATALI to retrieve the aerosol type from an extensive dataset, with ~~miscellaneous variable~~ quality and physical content.

3.2 EARLINET Mahalanobis distance-based typing algorithm ~~EMD~~

The EARLINET Mahalanobis distance-based typing algorithm is a method specifically developed for the use on the EARLINET database with a high level of flexibility in order to adapt to the different lidar set-ups and needs (Papagiannopoulos et al., 2018). The algorithm applies the Mahalanobis distance classifier (Mahalanobis, 1936) to classify observations into ~~aerosol classes.~~This maximum 8 (Dust, Volcanic, Mixed Dust, Polluted Dust, Clean Continental, Mixed Marine, Polluted Continental, Smoke) and minimum 4 (Dust, Maritime, PollutedSmoke, CleanContinental) aerosol classes, considering the needs of each

user and the provided number of the intensive properties. The method demonstrated to be high performing for on aerosol typing using optical properties measurements in previous studies (e.g., Burton et al., 2012; Russell et al., 2014; Hamill et al., 2016).

~~The typing algorithm consisted of two steps. The first step is the setup of the reference dataset and the number of the aerosol classes. The reference dataset consists of well-characterized EARLINET data. The second step is the classification of unclassified data. This is done by calculating classification of each aerosol layer is made by calculating the distance of an observation from already defined classes and the predefined reference classes and by attributing each observation to a specific class based on the minimum distance. For improving reliability of the output, a screening procedure is applied to the minimum distance. First, the minimum distance with value higher than a threshold is considered outlier and is discarded. This threshold depends on the degrees of freedom. Given that the overall predictive accuracy of a typing algorithm depends on the predefined reference classes, the choice of the appropriate reference dataset is a crucial parameter. So, a well-characterized EARLINET dataset of observations from 2008 to 2010 was used to define the aerosol classes (Pappalardo et al. 2013; Papagiannopoulos et al. 2016a, and Schwarz 2016). As this dataset did not include all the typical aerosol components presented over Europe (i.e., 4 for 3 degrees of freedom. Second, an extra screening criterion is applied to almost equal values for more than one aerosol class, when the values are below the assumed threshold. For this, each distance is associated with a probability and, then Continental, Marine, Mineral dust and dust mixtures, Smoke, Volcanic ash), the observations were enhanced with additional ones (already published in literature), contributing to a total of 69 layers (Pappalardo et al., 2013; Papagiannopoulos et al., 2016a).~~

In a next step, a sensitivity analysis was performed in order to identify the classifying properties that provide the most adequate information to better predict of the correct aerosol class. The intensive properties that held the most weight among others in the classification appeared to be the backscatter-related Ångström Exponent at 355nm and 1064nm ($BAE_{355/1064}$), the normalized probability is required to be over 50% lidar ratio at 532 nm (LR_{532}), and the ratio of the lidar ratios (LR_{532}/LR_{355}).

Consequently, the initial reference dataset was split into a training and a validation dataset, in order to evaluate the predictive accuracy of the automatic method. Therefore, the training dataset was used to calculate the classification functions and then each observation was validated using the training dataset. Additionally, the effect of the depolarization information to the training phase was also investigated. To complement the reference dataset in this context, general literature values for particle linear depolarization ratio at 532 nm were used. Generally, the depolarization ratio is a really important parameter that can strengthen the typing procedure in the presence of aspherical layers, but Papagiannopoulos et al. (2018) showed that for the 5 and 4 classes, the particle linear depolarization ratio became less important (in this case the highest weight in the classification corresponds to the lidar ratio at 532 nm).

~~The Mahalanobis algorithm has already been evaluated in the framework of the EARLINET ACTRIS campaign during the summer. Finally, the assessment of the predictive performance of the algorithm was tested on a testing dataset. For this purpose, EARLINET data collected during the ACTRIS Summer 2012 (Papagiannopoulos intensive measurements (Sicard et al., 2018) - Eight reference aerosol classes which are characteristic of the following general aerosol types were initially defined: clean continental, polluted continental, pure dust, mixed dust (Dust+Maritime), polluted dust (Dust+Smoke and/ or Dust+Polluted Continental), mixed maritime, smoke, and volcanic. The 2015; Granados-Muñoz et al., 2016b) were chosen to test the automatic~~

typing algorithm. The testing dataset comprised of 47 layers, 21 of which yielded depolarization ratio values. The performance of the algorithm was checked for each of the grouping classes (i.e., 8,7,6,5,4) and the predictive accuracy of the algorithm further-increased-increased up to 90% when the aerosol classes that tend to reflect the same optical properties value-values were combined into 4 (Dust, Maritime, Polluted Continental + Smoke, Clean Continental) The b-related Ångström Exponent at 355nm and 1064nm ($AE_{355/1064}$), the lidar ratio at 532 nm (LR_{532}), and the ratio of the lidar ratios (LR_{532}/LR_{355}) were used for the aerosol classification without providing the information of the depolarization. The study concluded that the fewer aerosol classes (i.e., 4, 5, 6 classes) could provide better-prediction-accuracy a successful prediction accuracy, even without depolarization values, but, nonetheless, a coarser and less insightful classification. Dust showed high prediction-classification showed a high success rate, whilst the aerosol types that performed worse were the smoke and polluted continental aerosol. However, when these two categories were combined into a single aerosol class, the correct prediction increased. More detailed description about the algorithm, the reference dataset and the set of intensive-parameters-the intensive parameters defined to separate different aerosol types can be found in Papagiannopoulos et al. (2018).

The output of the algorithm are: (i) the calculated Mahalanobis distance of the layers from each aerosol class and (ii) the calculated chi-square probability. For improving the reliability of the output, two screening criteria are applied to the calculated distances following the procedure of Burton et al. (2012). The first one has to do with the minimum accepted distance which depends on the number of degrees of freedom (i.e., the type dependent intensive properties) and the second one has to deal with the equal distances. In our study, we use the aforementioned 3 classifying parameters (i.e., the backscatter-related Ångström Exponent at 355nm and 1064nm ($BAE_{355/1064}$), the lidar ratio at 532 nm (LR_{532}), and the ratio of the lidar ratios (LR_{532}/LR_{355})) and the minimum accepted distance for a measurement to be typed is 4. When the distance is higher than the defined threshold, which means that no similarity with the aerosol classes of the reference datasets is found, the observation is not typed. So, the first filter ensures that the calculated distance has 99% probability (cumulative probability) to belong to the class that was estimated. If this happens for more than one classes, then the second filter is applied. According to this one, the normalized probability (the probability based on the calculated distances) of the aerosol class needs to be higher than 50% (Papagiannopoulos et al., 2018). Overall, the application of the two screening criteria, enhances the confidence of the classification. However, the selection of a strict threshold (i.e., normalized probability greater than 60%) enhances the correct prediction, but reduces the identification rate of the algorithm.

3.3 Methodology

Fifty four (54) Raman lidar cases of aerosol measurements (backscatter coefficient profiles at 1064nm, 532nm and 355nm, as well as the extinction coefficient profiles at 532nm and 355nm) over Thessaloniki during the period 2012-2015 were used for this study. These input parameters were processed with NATALI software-algorithm for the identification of the layer boundaries, the calculation of their mean intensive optical parameters and their corresponding uncertainties.

The NATALI typing was performed in the low resolution typing configuration (5 predominant aerosol types - Dust, Smoke, Continental Polluted, Continental and Maritime) since particle linear depolarization ratio measurements for Thessaloniki are were not available for the study period.

A sensitivity analysis of NATALI input parameters to

In what follows, we merged the output types from NATALI that tend to reflect the same aerosol characteristics, and hence we evaluate the corresponding effects on the typing rate of the algorithms. Thus, the smoke and the defined aerosol layers was conducted. Three different input configurations were evaluated, concerning the layer thickness (i.e., 300 m and 600 m), the smoothing parameters (i.e., 700 or 1000) and the threshold confidence (Minimum accepted confidence and Minimum agreement ratio 70% and 90%). From the processing with different settings we conclude that polluted continental categories were grouped into the more generic type of small particles with high lidar ratio values (PollutedSmoke). The selection of four main aerosol classes stems from the availability of intensive properties, the typing seemed stable independently to the different input settings. Specifically, when switching from Input Settings difficulty in deriving a confident classification without particle linear depolarization ratio and the difficulty in discriminating polluted continental and smoke particles that reveal the same type characteristics. Regardless, the aerosol classes describe the major aerosol components. The identified layer boundaries from NATALI are used as input in the EARLINET Mahalanobis distance-based typing algorithm. Considering the aforementioned typing merging, the EARLINET Mahalanobis distance-based typing algorithm was set to classify observations into 4 aerosol classes: CleanContinental, Dust, Maritime and PollutedSmoke.

Table 1 (i. e., filter window 700, layer depth 300 m and minimum accepted confidence 70%) to Input Settings 2 (i.e., filter window 1000, layer depth 600 m and minimum accepted confidence 70%), the number of layers identified is reduced, whereas by switching from Input Settings lists the aerosol types used in the aerosol classification and their correspondence with the available types provided by the two typing methods and Table 2 to Input Settings 3 lists the mean aerosol optical properties of the reference aerosol types defined by the two algorithms. The idea here is to compromise: i) the resolution (low) of the automatic classification owing to the availability of the optical properties (i.e., filter window 700, layer depth 600 m and minimum accepted confidence 90%) the percentage of the identified types changes, but the aerosol type remains the same. Layers classified as Unknown (24), when one or two intensive optical parameters were outside the acceptable limits and layers classified as N/A (3+2) were excluded from the analysis.

The lidar classification (lidar configuration), and ii) the type definition, which does include the wide spectrum of the aerosol types provided by the two automated typing techniques. As it is evident from the values presented in Table 2, the lidar classification scheme consists of the main classes: large (i) large particles with medium lidar ratios values particles (i.e., dust-like particles), large (ii) large particles with low lidar ratios value particles (i.e., marine-maritime particles), small (iii) small particles with high lidar ratios value particles (i.e., pollution and/or smoke particles) and small (iv) small particles with medium lidar ratios value particles (i.e., clean continental particles). The selection of four aerosol classes stems to the availability of intensive properties and the difficulty to derive a confident classification without particle linear depolarization ratio. Regardless, the aerosol classes describe the major aerosol components. Thus, both automated typing methods are set to classify observations into 4 aerosol classes: Clean Continental, Dust, Maritime, Polluted Smoke.

The EARLINET Mahalanobis distance-based typing algorithm classes are configured as follows: Dust=dust+volcanic+mixed dust+polluted dust, Polluted Smoke=smoke+polluted continental, Clean Continental=clean continental, and Maritime=mixed maritime. The identified layer boundaries from NATALI are used for input to the EARLINET Mahalanobis distance-based

typing algorithm typing algorithm. As, for NATALI, the low resolution typing consists of 5 aerosol classes, Smoke and Polluted Continental layers defined by NATALI were merged into Polluted Smoke (Polluted Continental + Smoke).

Desert dust layers (i.e., large with medium lidar ratio values particles) Generally, desert dust layers have optical properties that are considerably different from the other types, ~~thus they are easily identified~~. Their big size leads to low Ångström exponent values and the reported lidar ratio at 355nm ranges from 47 to 58 sr for Thessaloniki (Siomos et al., 2018), ~~with minimum values reported in spring in Thessaloniki~~. Maritime particles (i.e., large with low lidar ratio values), in contrast to desert dust particles yield low particle lidar ratio values. Polluted Smoke particles (i.e., small with high lidar ratio values particles) ~~),~~ PollutedSmoke particles are highly absorbing particles ~~exhibiting lidar ratios ranging from 52 to 73 sr (Amiridis et al., 2009; Baars et al., 2012)~~. The highest values appear during summer period in Thessaloniki (Siomos et al., 2018). Continental particles (i.e., small with medium lidar ratio values) present low, with high lidar ratio values, (i.e., 20–40 sr) and relatively high Ångström exponents (i.e., 1.0–2.5). ~~The categorization of continental layers~~. CleanContinental categorization is not completely straightforward, because the continental particles can be due attributed to different subcategories (i.e., local, continental polluted or mixtures). In general, the CleanContinental cases are typically elevated layers, i.e. layers not related to the local atmospheric boundary layer where the pollution and anthropogenic contribution would mean more absorbing particles and therefore labeled as PollutedSmoke aerosol.

4 Application of the two automatic algorithms to EARLINET data - case studies

~~The consistency~~

The application of the two automated aerosol typing algorithms on aerosol classification for ~~two cases was first examined~~ three cases is presented and discussed in detail. A Saharan dust outbreak ~~and~~, a biomass burning episode case and a complex aerosol structure scene over Thessaloniki station are presented below. The typing results are compared against manually typed profiles that were characterized via satellite using satellite data and model simulations. Specifically, in order to identify the source of aerosol particles, backward trajectories are calculated using the Hybrid Single-Particle Lagrangian Integrated Trajectory model (HYSPPLIT; Stein et al., 2015). Additionally, the BSC-DREAM8b model (e.g., Basart et al., 2012b) was used to verify the presence of Saharan dust. Finally, fire spots from the space-based MODIS sensor (<https://firms.modaps.eosdis.nasa.gov/map/>) product FIRMS (Fire Information for Resource Management System; Giglio et al., 2016) were used to confirm the presence of smoke over ~~of~~ Thessaloniki.

4.1 Dust Case

The first case refers to the occurrence of a dust plume over the station of Thessaloniki on 20th of May 2013. The retrieved profiles of the particle backscatter and extinction coefficient, lidar ratios, and Ångström exponents are shown in Fig. 1 (a-d). The measurement is characterized by two particle layers, the first one between 0.7 and 1.6km and the second one between 2.0 and 2.7km. The dust presence was confirmed by the ~~7-days~~ 4-days backward trajectories arriving at Thessaloniki on ~~20~~ 20th of May 2013 (Fig. 1e). The trajectories indicated an event of transported Saharan dust. Additionally, the BSC-DREAM8b

model was used to verify the presence of Saharan dust. The aerosol main layers and NATALI typing results are presented in Fig. 1f. For each aerosol layer (i.e., Layer **A1**: 0.7-1.6 km and Layer **B2**: 2-2.7 km), the mean optical properties were calculated using NATALI. The lidar ratio value for ~~layer A~~ Layer 1 is 45 ± 3 sr for 355nm and 43 ± 1.1 sr for 532nm and for ~~layer B~~ Layer 2 is 45 ± 1.4 sr for 355nm and 41 ± 1.3 sr for 532nm correspondingly. The ~~stability ratio~~ of the lidar ratio values (LR_{532}/LR_{355}) for Layer 1 is 0.94 ± 0.083 and for Layer 2 is 0.92 ± 0.13 . The mean ~~$AE_{355=532}$~~ $AE_{355/532}$ was 0.33 ± 0.06 for ~~layer A~~ Layer 1 and 0.38 ± 0.07 for ~~layer B~~ Layer 2 respectively. The mean ~~$BAE_{355=532}$~~ and mean $BAE_{532=1064}$ ~~$BAE_{355/532}$~~ mean $BAE_{532/1064}$ and mean ~~$BAE_{355/1064}$~~ mean $BAE_{355/1064}$ were 0.24 ± 0.06 ~~and~~ , 0.28 ± 0.04 ~~for the layer A~~ and 0.32 ± 0.22 for the first layer, while the values for the ~~layer B~~ Layer 2 were 0.13 ± 0.07 ~~and~~ , 0.06 ± 0.04 ~~and 0.05 ± 0.1~~ respectively. The ~~stability of the lidar ratio and Ångström exponent values could be considered as an indicator of homogeneity and small variability of the aerosol type within the layer.~~ The retrieved values indicate the presence of coarse particles and are in agreement with the typical dust values observed over Thessaloniki (Siomos et al., 2018). Also, Müller et al., (2007) reported values of mean backscatter-related Ångström exponent $BAE_{355/1064}$ of 0.2 ± 0.2 for desert dust events. NATALI typing output indicates dust layers (Fig. 1f). The EARLINET Mahalanobis distance-based typing algorithm also classifies ~~layers A and B~~ Layers 1 and 2, as dust ~~ease cases~~ (Mahalanobis distance is minimum), and Mahalanobis probability also shows good predictive performance (68% ~~layer A~~ Layer 1 and 44.6% ~~layer B~~ Layer 2). (Fig. 1(g-h)). The plotted probabilities are assigned to each calculated distance and are different from the normalized ones that used as a screening criteria (and add up to 100).

4.2 Biomass burning ~~episodic case~~ episode case

The second case is a biomass burning episode that occurred on 2nd September 2013. The ~~3-day~~ backward trajectories from HYSPLIT in conjunction with fire spots from MODIS satellite product FIRMS indicate the biomass burning episode, transported from central Europe in the region of Thessaloniki. The optical profiles and the layers of aerosols are shown in Figure 2 (a-d). The measurement is characterized by three particle layers: the first one is between 0.98km and 1.2km (Layer **A1**), the second one between 1.7 and 2.6km (Layer **B2**) and the third one between 2.9 and 3.5 km (Layer **C-3**). The upper and lower boundary of each layer is marked with lines. NATALI typing is presented in Fig. 2f. The mean values of all optical parameters for each detected layer were calculated. The mean Lidar ratio was calculated 69 ± 4 for 355nm and 70 ± 4 for 532nm (Layer **A1**), for ~~layer B~~ Layer 2 was 68 ± 2 sr for 355nm and 69 ± 2 sr for 532nm and Layer **C-3** has values of ~~6669 ± 1.4~~ 2.3 sr for 355nm and ~~7275 ± 1.5~~ 2.5 sr for 532nm. The ratio of the lidar ratios values (LR_{532}/LR_{355}) for Layer 1 is 1.02 ± 0.004 , for Layer 2 is 1.02 ± 0.007 and for Layer 3 is 1.09 ± 0.048 . Mean Ångström (AE) exponent for each layer was estimated at 1.99 ± 0.13 (~~layer A~~ Layer 1), 1.84 ± 0.06 (~~layer B~~ Layer 2) and 1.61 ± 0.05 (~~layer C~~) 0.08 (Layer 3). These values are consistent with previous reported ones, e.g., Tesche et al. (2011a) reported values of the extinction Ångström exponent for smoke of 0.9 ± 0.26 . The mean backscatter-related Ångström exponent (BAE) at 355-532nm was 2.03 ± 0.13 (~~layer A~~ Layer 1), 1.88 ± 0.06 (~~layer B~~) and 1.79 Layer 2) and 1.80 ± 0.05 (~~layer C~~) 0.08 (Layer 3), while the mean backscatter-related Ångström exponent (BAE) at 532-1064nm was 1.43 ± 0.07 (~~layer A~~ Layer 1), 1.08 ± 0.04 (~~layer B~~) and 1.18 Layer 2) and 1.03 ± 0.05 (Layer 3), and mean backscatter-related Ångström exponent (BAE) at 355-1064nm was 1.63 ± 0.07 (Layer 1), 1.36 ± 0.04 (Layer 2) and 1.31 ± 0.03

(~~layer C~~Layer 3), respectively. These values are in accordance with the typical biomass burning values observed over Thessaloniki. ~~(Giannakaki et al. (2010) reported an annual mean lidar ratio at 355nm of 69 ± 17 sr and a mean BAE at 355-532nm of 1.7 ± 0.7 , while Siomos et al. (2018) (2018a) found lidar ratio at 355nm ranging from 51 to 73 sr for biomass burning events. NATALI aerosol output (Fig. 2f) and the EARLINET Mahalanobis distance-based typing and the probability values algorithm values (distance and probability) also confirm the presence of smoke layers (Fig. 2(g-k)). The plotted probabilities are assigned to each calculated distance and are different from the normalized ones that used as a screening criteria (and add up to 100).~~

4.3 Aerosol classification using NATALI Mixed case

The third case is a case with complex aerosol structure that occurred on 21st August 2014. This case offers the opportunity to check the reliability of the algorithms in conditions where different aerosol types in different layers exist. The 5-day backward simulation from HYSPLIT indicates the pattern of the origin of air masses before reaching the study area as presented in Figure 3e. The motion of the particles is in a southwest direction from a dust source towards Thessaloniki. So, the path of the air masses arriving over Thessaloniki suggests a mixture of dust (at height of about 3 km) and continental and marine particles. The optical profiles and the layers of aerosols are shown in Figure 2 (a-d). The measurement is characterized by four particle layers: the first one is between 0.9km and 1.10km (Layer 1), the second one between 1.5 and EARLINET Mahalanobis 2.1km (Layer 2), the third one between 2.5 and 3.8 km (Layer 3) and the fourth one between 4.47 and 5.52km. The upper and lower boundary of each layer is marked with lines. NATALI typing is presented in Fig. 2f. The mean values of all optical parameters for each detected layer were calculated. The mean Lidar ratio was calculated 56 ± 3.17 for 355nm and 37 ± 2.09 for 532nm (Layer 1), for Layer 2 was 57 ± 1.75 sr for 355nm and 44 ± 1.35 sr for 532nm, Layer 3 has values of 43 ± 0.91 sr for 355nm and 42 ± 0.68 sr for 532nm and Layer 4 has values of 29 ± 0.77 sr for 355nm and 30 ± 0.91 sr for 532nm. The ratio of the lidar ratios values (LR_{532}/LR_{355}) for Layer 1 is 0.66 ± 0.08 , for Layer 2 is 0.77 ± 0.05 , for Layer 3 is 0.97 ± 0.067 and for Layer 4 is 1.03 ± 0.1 . Mean Ångström exponent (AE) for each layer was estimated at 2 ± 0.14 (Layer 1), 0.86 ± 0.08 (Layer 2), 0.85 ± 0.05 (Layer 3) and 1.72 ± 0.07 (Layer 4). The mean backscatter-related Ångström exponent (BAE) at 355-532nm was 0.98 ± 0.14 (Layer 1), 0.19 ± 0.08 (Layer 2), 0.10 ± 0.05 (Layer 3) and 1.21 ± 0.06 (Layer 4). The mean backscatter-related Ångström exponent (BAE) at 532-1064nm was 0.77 ± 0.08 (Layer 1), 0.68 ± 0.04 (Layer 2), 0.58 ± 0.03 (Layer 3) and 1.9 ± 0.03 (Layer 4). The mean backscatter-related Ångström exponent (BAE) at 355-1064nm was 0.78 ± 0.07 (Layer 1), 0.49 ± 0.04 (Layer 2), 0.39 ± 0.03 (Layer 3) and 1.64 ± 0.03 (Layer 4), respectively. NATALI aerosol output is presented in Fig. 2f. NATALI classifies as Unknown the Layer 1 and Layer 2, as the ANNs could not identify any aerosol type (no ANNs passed the threshold criteria), whilst classifies Layer 3 as Dust and Layer 4 as Maritime. The EARLINET Mahalanobis distance-based typing algorithm typing probability values are presented in Fig. 2 (g-j). The Layer 1 is classified as CleanContinental, the Layer 2 and the Layer 3 as Dust and the Layer 4 as CleanContinental. Both typing techniques classify Layer 3 as Dust. The values of the optical parameter indicate the presence of coarse particles and are in agreement with the typical dust values observed over Thessaloniki (Siomos et al., 2018). Layer 4 is typed differently. This layer is recognized by NATALI as Maritime one (lower Lidar Ratio values), while in the EARLINET Mahalanobis distance-based typing method is attributed as signature for CleanContinental. Layer 1 and Layer 2 are typed as Unknown layers from NATALI. However, Layer 1 is recognised as CleanContinental by EARLINET Mahalanobis

distance-based typing method, with mean backscatter-related Ångström exponent and Lidar Ratio values, representative of the continental particles, allowed in the EARLINET Mahalanobis distance-based typing method (Table 2). Given that, the two algorithms take into consideration different combination of the intensive optical properties (i.e., NATALI uses AE), this is also a reason that differentiate the output retrieval. The layer high Ångström exponent of Layer 1, considered in the NATALI typing algorithm, maybe is the reason that ANNs answers are filtered out. Same for Layer 2, NATALI return the layer as Unknown, due to the low confidence level of the ANN outputs. However, the EARLINET Mahalanobis distance-based typing method classifies the Layer as Dust, based on the values of the optical parameters used for typing (Table 2). This example demonstrates that type assignment is quite challenging and not always possible in ambiguous atmospheric scenes.

4.4 Automated aerosol classification using NATALI and EARLINET Mahalanobis distance-based typing algorithm

The complete Thessaloniki multiwavelength Raman lidar dataset for the 2012-2015 period was analyzed in terms of aerosol typing with the NATALI and the EARLINET Mahalanobis distance-based typing algorithm using the methodology-automated methodologies reported in Sect. 3.33.1 and 3.2. For the 54 cases, 116-120 layers were identified by the layer identification module of NATALI. Out of 116 layers, NATALI classified 80 layers, 26 were flagged as Unknown or N/A, and, out of which 10 layers were rejected due to the fact that they were below 1 km. On the other hand, the (below overlap region).

NATALI classified 82 of the total 110 layers (75% of the total number of cases) and 28 were flagged as Unknown layers. These missed cases are based mostly on retrievals with large uncertainties of the optical parameters and the calculated uncertainties and rely on the minimum acceptance confidence that not allowed ANNs to return any aerosol type. However, these 'unknowns' layers provided type results (93%), when the minimum acceptance confidence was set to 70%. However, in this study we investigate the application of NATALI algorithm holding the retrievals that pass a higher percentage of quality checks. The EARLINET Mahalanobis distance-based typing algorithm successfully classified 114 layers-classified 98 layers (89%) and 12 layers were not assigned to any cluster (the Mahalanobis distance values larger than 4), showing a higher identification rate. However, the comparison is made to the 80 layers that were identified by both algorithms; this rate also depends on the screening criteria. When we set the second of the screening criteria (i.e., the normalized probability of the aerosol class) to be higher than 50%, the number of untyped layers increased (22 untyped cases, 80% typing rate). Overall, the identification rate can be changed regarding the needs of the user and the selection of the appropriate threshold criteria. In addition, we examined the ability of the both methods to provide typing results, relative to the derived layer optical depth (Aerosol Optical Depth - AOD) values. This study showed that the mean layer AOD of the typed layers has similar value for both typing techniques (0.1076 ± 0.0898 (SD) for NATALI and 0.11166 ± 0.0928 (SD) for the EARLINET Mahalanobis distance typing algorithm). However, the EARLINET Mahalanobis distance algorithm seems more sensitive to lower values (not typed cases show a mean layer AOD of 0.0298 ± 0.0578 for the EARLINET Mahalanobis distance algorithm, in contrast to the 0.06899 ± 0.0606 for NATALI). These low AOD values are associated with higher uncertainties on the intensive properties.

The two algorithms attributed the same aerosol type to 58 layers, showing an agreement of 72.5% of the typed cases. In Fig. 3, the number of detected layers per typing class are presented for both algorithms for the whole study period. The typing procedures show the predominance of PollutedSmoke category for Thessaloniki, followed by the CleanContinental category.

~~Dust results in about~~ Figure 4 (a,c) presents the percentages of the total typed layers into the defined aerosol classes for both automated algorithms. Also, the untyped layers from the EARLINET Mahalanobis distance-based typing algorithm that were typed by NATALI and visa versa are presented in Fig. 4 (b,d). The EARLINET Mahalanobis distance-based typing algorithm, classified 27 of the 28 untyped layers by NATALI, as follows: 6 layers as Dust (21.4%), ~~10 % of the observed layers~~. Finally, ~~although Thessaloniki is a coastal site, the maritime layers are rare, due to the mixing with other aerosol types. These results are in agreement with the findings of Siomos et al., 2018.~~

as PollutedSmoke (35.7%), 10 as CleanContinental (35.7%) and only 1 layer as Maritime (3.6%). Here again, the EARLINET Mahalanobis distance-based typing algorithm, indicated PollutedSmoke and CleanContinental category as the main sources of the aerosol layers observed, while Maritime category is the less pronounced one. On the other hand, NATALI classified 11 of the ~~12 untyped layers from the EARLINET Mahalanobis distance-based typing algorithm, as follows: 2 layers were typed as Dust (16.7%), 6 as PollutedSmoke (50%) and 3 layers as Maritime (25%). In general, NATALI software seems more sensitive to the Maritime category. In particular, the agreement is high enough reasonably close for the desert dust cases (12.5% and 1510% and 17% for NATALI and EARLINET Mahalanobis distance-based typing algorithm, respectively).~~ ~~A larger difference is observed, nevertheless, it becomes evident that the particle linear depolarization ratio could increase the ability for correctly predicting dust particles (Fig. 4). The agreement is remarkable for the PollutedSmoke and CleanContinental with 15-17 cases of mismatch, with a higher occurrence of PollutedSmoke cases for the cases, with occurrence ratio of 43% and 47% for NATALI and EARLINET Mahalanobis distance-based typing algorithm and of CleanContinental for the NATALI algorithm., respectively, whilst the CleanContinental cases showed a difference with 18% and 24% for NATALI and EARLINET Mahalanobis distance-based typing algorithm, respectively. Also, difference is observed for the Maritime category, where one case was attributed to this category by NATALI.~~

The differences observed in Fig. 34, primarily, can be described by the different definition of the aerosol classes for the two typing methods. It is worth noting that the aerosol model for the NATALI typing method is trained based on synthetic data. The comparison between the aerosol model values used for the training and the reported literature values showed differences (Nicolae et al., 2018), as follows: Synthetic values found lower than those observed for continental-rural (AE 350/550), continental-polluted (CR 500/1000), and dust (CR 500/1000) types and synthetic values are greater for continental-rural (LR350) and volcanic (DEP 550) types. The reasons for these differences could be related on one side to uncertainty or variability issues of the measured values and on the other side to the aerosol model limitations e.g. due to spheroidal model and mono-modal log-normal distribution (Nicolae et al., 2018). On the other hand, for the EARLINET Mahalanobis distance-based typing method the aerosol reference class is defined using an ensemble of high quality EARLINET observations. However, different measurement locations may be a reason for mismatches in the attributed aerosol category. Moreover, the ~~different approach by the typing general different approaches of the two methods may have an impact to on the aerosol typing.~~ ~~For instance, for the EARLINET Mahalanobis distance-based typing algorithm the quality of the reference data and, even that both classifications schemes have been trained using supervised learning techniques. Additional to the ability of the two automated algorithms to attribute the correct type to the aerosol layers, also the uncertainties of the input measurements have to be taken~~

into account. Both algorithms are unable to return aerosol type for not perfect calibrated input data, i.e., when one or more intensive optical properties have large uncertainties.

In what follows, a comparison is made to the 71 common layers, identified by both algorithms. Figure 5 presents the number of detected layers per typing class attributed by each algorithm for the whole study period. The two algorithms attributed the same aerosol type to 56 layers, showing an agreement of 78.8% of the typed cases. The typing procedures show the predominance of the PollutedSmoke category for Thessaloniki, followed by the CleanContinental category. This result is compliant with the CALIPSO scheme at global level and with results about Thessaloniki site characterization done by Siomos et al. (2018b). In their paper Siomos et al. (2018b), used data from a double monochromator Brewer spectrophotometer and a sunphotometer in order to classify aerosol cases during the period 2007-2017 in Thessaloniki: they found that the Water Soluble category (which can be related to the Clean Continental consisting of water soluble particles) correspond to 29.1% of the cases, which is in fair agreement with results reported in Fig. 5. Dust results in a smaller amount of the observed layers, according to the retrievals of both algorithms. Finally, although Thessaloniki is a coastal site, the correct definition of aerosol classes are of paramount important. It is worth noting that the aerosol model for the NATALI typing method is based on synthetic data, whereas for the EARLINET Mahalanobis distance-based typing method the aerosol model is based on EARLINET observations maritime layers are rare, presumably, due to the mixing with other aerosol types. These results are in agreement with the findings of Siomos et al. (2018).

Although, each automated classification algorithm has important differences acknowledged above, the comparison showed an overall good agreement, especially for dust particles. This happens because for the four defined aerosol classes. The convergence of the two different methods on the same type can be regarded as a signature of reliability. An almost perfect score was found for the PollutedSmoke category, following by the Dust one, given that the dust class is very well defined for both typing schemes and, as the physical properties are different of dust particles differentiate from the other three classes. Considering the mismatches in the CleanContinental category, good agreement is found. By contrast, the maritime category is defined in a different way for the two automated algorithms. The EARLINET Mahalanobis distance-based typing algorithm considers maritime layers mixed with other aerosol types, whereas for the NATALI the mixing is negligible and the aerosol type refers to pure maritime aerosol. Therefore, a direct comparison of the two methods is not possible. The absence of measurements for such kind of particle also did not allowed a direct assessment of the pure marine particle synthetic data into NATALI algorithm itself. The cases case typed as Maritime by NATALI were identified as Dust or was identified as CleanContinental: this is because of the different lidar ratio at 532nm and backscatter Ångström related values allowed in the NATALI scheme which are recognized by the EARLINET Mahalanobis distance-based typing method as signature for Dust or CleanContinental types (Table 1). Finally, the largest number of mismatches is related to the PollutedSmoke and the CleanContinental categories.

Discussion of the mismatch cases is extremely important for understanding the limitations of the assumptions behind the two schemes and provide a useful data set to investigate the causes, which will help to potentially improve the typing methods. In what follows, the NATALI aerosol model dataset is ingested into the EARLINET Mahalanobis distance-based typing algorithm with the aim of understanding the differences of in the aerosol definition. 259407 out of 424236 instances (61%) were typed by

the EARLINET Mahalanobis distance-based typing algorithm. The not-typed instances were discarded by the algorithm owing to a distance-threshold filter and a distance-similarity filter (see Sect. 3.2). The results of the application of the ~~Mahalanobis EARLINET EARLINET Mahalanobis~~ distance-based method to the modeled reference classes of NATALI are presented in Fig. 4–6. Each percentage is defined here as follows: The denominator in each column is the number of NATALI samples for a given type, and the numerator is the number of the EARLINET Mahalanobis distance-based classifications for the same sample.

Good-Fair agreement is found for the NATALI synthetic datasets of Dust and PollutedSmoke categories. In particular, for Dust reached agreement of 83.7% (100166 typed over 119650). This result is very good considering that up to 30% of the aerosol class contains mixtures. For the ~~Polluted-Smoke~~PollutedSmoke, the majority (58.7%) of the synthetic data, corresponding to 48180 instances from a total typed of 82041 (58.7%) is classified as PollutedSmoke. For the CleanContinental type, the agreement is poor (15%), as only 7605 over 52058 were correctly classified and the majority of the instances (36552) is typed as Dust (70%). This Dust-CleanContinental mismatch is confirmed also by the 11.8% of the NATALI reference dust cases that are typed as Clean Continental by the EARLINET Mahalanobis distance-based typing method. ~~Furthermore, 15%~~ Furthermore, 7901 of the instances (15%) are classified as PollutedSmoke. Finally, the agreement for Maritime is 23% of the cases (1319 from a total of 5658 cases), indicating the different class ~~definition-definitions~~ as well as the big spread of the NATALI maritime model that ~~tends-causes~~ the EARLINET Mahalanobis distance-based typing algorithm to misclassify those data.

Despite the important differences found on the set-up of the two methods, a very good agreement is achieved when the algorithms are applied on the EARLINET Thessaloniki dataset as shown in Fig. 5. ~~For Dust and Polluted-Smoke 7. For~~ For PollutedSmoke and CleanContinental the accuracy reached ~~88.8% and 93.7% respectively. Very 88%~~ (the common number of typed layers as PollutedSmoke from both algorithms to the total layers typed as PollutedSmoke from NATALI) and 65% (the common number of typed layers as CleanContinental from both algorithms to the total layers typed as CleanContinental from NATALI) respectively. A good performance is also found for the ~~Clean-Continental category (70)~~Dust category (78%). It is obvious, that the additional information of the particle depolarization ratio will further improve the performance of the automated typing algorithms in distinguishing dust particles from other aerosol types. A not satisfactory agreement is observed for the Maritime, which is the aerosol type less encountered over Thessaloniki. ~~It is worth mentioning, here, that the addition of the particle depolarization ratio information will improve the performance of the automated typing algorithms and offer a more detailed aerosol classification~~Overall, there were 15 cases (on Thessaloniki dataset) that the two methods provided different typing results. There are seven cases typed as CleanContinental aerosol by NATALI and PollutedSmoke or CleanContinental by EMD, five cases typed as PollutedSmoke by NATALI and Dust by EMD, 2 cases typed as Dust by NATALI and CleanContinental by EMD and one case type as Maritime by NATALI and CleanContinental by EMD. These mismatches are illustrated in Figure 7. In order to understand these differences we highlight some critical issues relevant to the type definitions of the two methods, based on Nicolae et al., (2018) and Papagiannopoulos et al., (2018):

CleanContinental: The contribution of the Soot component in the chemical composition of the cleancontinental category, allows higher lidar ratio values at 532nm (52–53sr) in the NATALI scheme. Consequently, layers recognized by NATALI as CleanContinental ones, in the EARLINET Mahalanobis distance-based typing method are attributed as signature for PollutedSmoke or Dust (Figure 7).

5

Marine: As observations of pure maritime particles are quite scarce within EARLINET and, generally, when these particles are observed their characteristics are far from pristine, the Maritime category for the EARLINET Mahalanobis distance typing algorithm, corresponds to mixed maritime layers. This is different from the pure maritime category that NATALI identifies. Lower lidar ratio values for 532nm (19–25sr) are defined in the NATALI software for the identification for marine layers, in contrast to the higher ones (16–32sr) allowed in the EARLINET Mahalanobis distance typing algorithm.

10

Dust: Higher values of Lidar ratio at 532nm are allowed for EARLINET Mahalanobis distance typing algorithm in identifying dust particles (41–63sr), considering all dust-like aerosol types as one category, while NATALI allows values corresponding to more pure cases (44–49sr). Therefore, a number of Dust recognition are attributed to PollutedSmoke particles (6 cases) or CleanContinental particles (1 case) in the NATALI output.

15

PollutedSmoke: An almost perfect score was found for the PollutedSmoke category. This can be attributed to the similar reference values attributed by both typing algorithms. The values of Lidar ratio at 532nm allowed for NATALI are in the range of 62–92sr (ContinentalPolluted and Smoke) and the ones allowed in the EARLINET Mahalanobis distance typing algorithm are in the range of 50–97sr.

20

As discussed above, the ability and performance of both typing methods strongly depend on the differences on the reference aerosol classes definition and the selection of the applied thresholds. The latter follow a different approach in the typing schemes (are linked with uncertainties in the NATALI algorithm and with the normalized probability based on the calculated distanses in the Mahalanobis distance typing algorithm).

25

5 Summary and Conclusions

In this study, two automated typing methods ~~were are~~ used to obtain the ~~best estimate of the~~ dominant aerosol type using optical properties, measured by a multiwavelength Raman lidar over Thessaloniki. The ~~prediction application~~ of the automatic classification methods in ~~two study cases showed positive~~ three case studies shows consistent results when compared against manually classified EARLINET data. Additionally, all Raman measurements for the period 2012 to 2015 ~~were are~~ analyzed and a coarse aerosol classification analysis ~~was made and is made~~. The classification analysis covers the four major aerosol classes ~~were defined as follows: 1) Dust, 2) Maritime, 3) Polluted Smoke and 4) Clean Continental.~~

30

Sensitivity processing in the settings of NATALI layering modules did not affect the agreement of the two algorithms which is similar (66%, 65% and 70% for each input setting correspondingly), only changes in the number of identified layers being observed. Good agreement in the aerosol classification was obtained between the two automatic procedures despite their different approaches, representing conditions in the area of Thessaloniki, i.e., Dust, Maritime, PollutedSmoke and CleanContinental.

Both algorithms indicated

Both classification schemes indicate PollutedSmoke and CleanContinental category as the main sources of the aerosol layers observed in 3 year-years of measurements over Thessaloniki, while Dust accounted for about 12-15% accounts for about 10-17% of the cases and Maritime particles were rarely observed. Good Remarkable agreement is found for the Dust and Polluted Smoke (83.7% and 58.7%, respectively). Low in the aerosol typing for the PollutedSmoke category (88%), followed by Dust (78%), whilst fair agreement is observed for the Clean Continental and Maritime (15% and 23%, respectively), this CleanContinental category (65%). Maritime is the worst-performing category, underlining the need for further investigation on the identification and validation of optical property thresholds to identify sea salt aerosol. The convergence of the two different methods on the same type can be regarded as a signature of reliability, while differences can be ascribed to the different aerosol type definition. For instance, the NATALI algorithm considers pure marine layers while the EARLINET Mahalanobis distance-based typing algorithm considers a mixture of marine with other aerosol types.

The comparison of the typing results showed that the two algorithms are in a very good agreement of 88.8% for Dust, 93.7% for Polluted Smoke and 70% for Clean Continental. A poorer agreement is found for the Maritime type underlining the need for further investigation, which is partially valid also for Clean Continental.

In addition Moreover, despite the differences in the optical property thresholds allowed in each typing scheme, many of the schemes' thresholds do have overlap (especially on the low resolution mode), making it easy to affirm inferred aerosol type into multiple categories. In general, it is worth mentioning that, as demonstrated in Nicolae et al. (2018) and Papagiannopoulos et al. (2018), the availability of the particle linear depolarization ratio improves the predictive accuracy of both methods. Its availability could enhance the strength of correct predictions and, especially for the dust particles recognition and could lead to the increase of the number of detected types (high resolution typing) for both algorithms. It should also be mentioned that the success rate of the aerosol classification schemes (and thus their comparability) is largely dependent on the uncertainties of the optical properties and the threshold criteria defined. Future analysis comparing the two typing schemes on the high resolution mode, as well as further application on other lidar stations with more complex locations and topography, is ongoing work for determining the validity of the comparison made in this study.

Overall, we conclude that the application of these typing techniques could contribute to the improvement of the model treatment and to a better exploitation of the present and future data from satellite remote sensing for aerosol classification schemes. Further investigation on other EARLINET lidar stations needs to be conducted. Assessments about the adaptability of the presented aerosol type classification schemes to future satellite missions are ongoing work. Considering that up to now, there is not a well established method to provide correct results under every conditions, the systematic and full comparison of different methods is the only way to have a better insight about the possibility of inferring the aerosol source from the

observed optical properties. Additionally, the modularity of the EARLINET Single Calculus Chain tool offers the possibility to implement aerosol typing procedures for automatically providing the aerosol typing information. A comparison between the two existing typing schemes that make use of the aerosol optical properties could be rather useful, as it will allow from one side the optimization of aerosol typing models and on the other side the definition of the aerosol types used and the selection of the reference dataset.

Code availability. The lidar data used in this study are available at <http://data.earlinet.org>. NATALI algorithm is publicly available at <http://natali.inoe.ro/resources.html/software>. The EARLINET distance-base typing algorithm and related reference datasets will be soon available at www.earlinet.org.

Author contributions. D. Balis, N. Siomos, and KA. Voudouri performed and processed lidar measurements during the period 2012–2015. N. Papagiannopoulos developed the EARLINET Mahalanobis distance-based typing algorithm. D. Nicolae developed the NATALI aerosol typing algorithm. K. Michailidis carried out the processing of lidar measurements with the EARLINET Mahalanobis distance-based typing algorithm. KA. Voudouri carried out the processing of lidar measurements with NATALI algorithm and prepared the figures. L. Mona reviewed parts of the comparison of the two algorithms. D. Balis is the PI of the lidar station and directed the preparation of the manuscript. KA. Voudouri prepared the manuscript with contributions from all co-authors.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The research leading to these results has received funding from the ~~European Union's Horizon 2020 Research and Innovation programme under grant agreement no. 602014, ECARS (East European Centre for Atmospheric Remote Sensing) project, under grant agreement n. 654109, ACTRIS-2 project, and from~~ the European Union's Horizon 2020 Research and Innovation programme for Societal challenges - smart, green and integrated transport, under grant agreement no. 723986, ~~EUNADICS-AV project. Further financial support is provided by the Romanian Ministry of Research and Innovation throughout the Core National Program no. 33N/PN2018~~ (project EUNADICS -AV – European Natural Disaster Coordination and Information System for Aviation). Voudouri K.A acknowledges the support of the General Secretariat for Research and Technology (GSRT) and the Hellenic Foundation for Research and Innovation (HFRI), ~~no~~ Scholarship code: 95041. The financial support for EARLINET in the ACTRIS Research Infrastructure Project by the European Union's Horizon 2020 research and innovation program under grant agreement no. 654169 in the Seventh Framework Programme (FP7/2007–2013) and the ECARS project (East European Centre for Atmospheric Remote Sensing), under grant agreement n. 602014, ACTRIS-2 project, are gratefully acknowledged. Further financial support is provided by the Romanian Ministry of Research and Innovation throughout the Core National Program no. 33N/PN2018. ~~654109.~~

References

- Amiridis, V., Balis, D. S., Kazadzis, S., Bais, A., Giannakaki, E., Papayannis, A., and Zerefos, C.: Four-year aerosol observations with a Raman lidar at Thessaloniki, Greece, in the framework of European Aerosol Research Lidar Network (EARLINET), *Journal of Geophysical Research: Atmospheres*, 110, <https://doi.org/10.1029/2005JD006190>, d21203, 2005
- 5 Amiridis, V., 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, *Atmos. Chem. Phys.*, 9, 2431–2440, doi:10.5194/acp-9-2431-2009, 2009.
- [Amiridis, V., Marinou, E., Tsekeri, A., Wandinger, U., Schwarz, A., Giannakaki, E., Mamouri, R., Kokkalis, P., Biniotoglou, I., Solomos, S., Herekakis, T., Kazadzis, S., Gerasopoulos, E., Proestakis, E., Kottas, M., Balis, D., Papayannis, A., Kontoes, C., Kourtidis, K., Papagiannopoulos, N., Mona, L., Pappalardo, G., Le Rille, O., and Ansmann, A.: LIVAS: a 3-D multi-wavelength aerosol/cloud database based on CALIPSO and EARLINET, *Atmospheric Chemistry and Physics*, 15, 7127–7153, doi:10.5194/acp-15-7127-2015, <http://www.atmos-chem-phys.net/15/7127/2015/>, 2015.](https://doi.org/10.5194/acp-15-7127-2015)
- [Ansmann, A., Wagner, F., Müller, D., Althausen, D., Herber, A., von Hoyningen-Huene, W., and Wandinger, U.: European pollution outbreaks during ACE 2: Optical particle properties inferred from multiwavelength lidar and star-Sun photometry, *J. Geophys. Res.*, 107, 4259, <https://doi.org/10.1029/2001JD001109>, 2002.](https://doi.org/10.1029/2001JD001109)
- 15 Ansmann A., Bosenberg J., Chaikovskiy A., Cameron A., Eckhardt S., Eixmann R., Freudenthaler V., Ginoux P., Komguem P., Linne H., Angel Lopez Marquez M., Matthias V., Mattis I., Mitev V., Müller D., Music S., Nickovic S., Pelon J., Sauvage L., Sobolewsky P., Srivastava M. K., 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), 4783, doi:10.1029/2003JD003757, 2003
- 20 Ansmann, A., Mattis, I., Müller, D., Wandinger, U., Radlach, M. and co-authors: Ice formation in Saharan dust over central Europe observed with temperature/humidity/aerosol Raman lidar. *J. Geophys. Res.* 110, doi:10.1029/2005JD005000, 2005.
- Ansmann, A., Petzold, A., Kandler, K., Tegen, I., Wendisch, M., Müller, D., Weinzierl, B., Müller, T., and Heintzenberg, J.: Saharan Mineral Dust Experiments SAMUM-1 and SAMUM-2: what have we learned?, *Tellus B*, 63, doi:10.1111/j.1600-0889.2011.00555.x, 2011.
- 25 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 PollyNET: 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, <http://www.atmos-chem-phys.net/16/5111/2016/>, 2016.
- 30 Balis D., Amiridis V., Zerefos C., Gerasopoulos E., Andreae M., Zanis P., Kazantzidis A., Kazadzis S., and Papayannis A.: Raman lidar and sun photometric measurements of aerosol optical properties over Thessaloniki, Greece during a biomass burning episode, *Atmospheric Environment*, Volume 37, Issue 32, 2003, Pages 4529-4538, ISSN 1352-2310, [https://doi.org/10.1016/S1352-2310\(03\)00581-8](https://doi.org/10.1016/S1352-2310(03)00581-8), 2003.

- Balis D., Amiridis V., Nickovic S., Papayannis A., and Zerefos C.: Optical properties of Saharan dust layers as detected by a Raman lidar at Thessaloniki, Greece, *Geophysical Research Letters - GEOPHYS RES LETT.* 311. 10.1029/2004GL019881, 2004.
- Basart, S., Pérez, C., Nickovic, S., Cuevas, E., and Baldasano, J.: Development and evaluation of the BSC-DREAM8b dust regional model over Northern Africa, the Mediterranean and the Middle East, *Tellus B*, 64, 18539, <https://doi.org/10.3402/tellusb.v64i0.18539>, 2012b.
- 5 Belegante, L., Talianu, C., Nemuc, A., and Nicolae, D.: Detection of local weather events from multiwavelength lidar measurements during the EARLI09 campaign, *Rom. J. Phys.*, 56(3–4), 484–494, 2011.
- Belegante, O., Nicolae, D., Nemuc, A., Talianu, C. and Derognat, C., Retrieval of the boundary layer height from active and passive remote sensors. Comparison with a NWP model, *Acta Geophysica*, 62 (2), 276-289, doi:10.2478/s11600-013-0167-4, 2014.
- Biniotoglou, I., Basart, S., Alados-Arboledas, L., Amiridis, V., Argyrouli, A., Baars, H., Baldasano, J. M., Balis, D., Belegante, L., Bravo-
10 Aranda, J. A., Burlizzi, P., Carrasco, V., Chaikovskiy, A., Comerón, A., D’Amico, G., Filioglou, M., Granados-Muñoz, M. J., Guerrero-Rascado, J. L., Ilic, L., Kokkalis, P., Maurizi, A., Mona, L., Monti, F., Muñoz-Porcar, C., Nicolae, D., Papayannis, A., Pappalardo, G., Pejanovic, G., Pereira, S. N., Perrone, M. R., Pietruczuk, A., Posyniak, M., Rocadenbosch, F., Rodríguez-Gómez, A., Sicard, M., Siomos, N., Szkop, A., Terradellas, E., Tsekiri, 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,
15 <https://doi.org/10.5194/amt-8-3577-2015>, 2015.
- Bosenberg, J., Matthias, V., Amodeo, A., et al.: EARLINET: A European Aerosol Research Lidar Network to Establish an Aerosol Climatology’, Report of the Max-Planck-Institute for Meteorology No. 348, 16–39, 2003.
- 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, *Atmospheric Measurement Techniques*, 5, 73–98, doi:10.5194/amt-5-73-2012, <http://www.atmos-meas-tech.net/5/73/2012/>, 2012.
- [D’Amico, G., Amodeo, A., Baars, H., Biniotoglou, 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, <https://doi.org/10.5194/amt-8-4891-2015>, 2015.](https://doi.org/10.5194/amt-8-4891-2015)
- [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. M., N. G., and Seefeldner: Depolarization ratio profiling at several wavelengths in pure Saharan dust during SAMUM 2006, *Tellus B*, 61, 165–179, <https://doi.org/10.1111/j.1600-0889.2008.00396.x>, 2009.](https://doi.org/10.1111/j.1600-0889.2008.00396.x)
- [Freudenthaler, V., Linné, H., Chaikovskiy, A., Rabus, D., and Groß, S.: EARLINET lidar quality assurance tools, *Atmos. Meas. Tech. Discuss.*, <https://doi.org/10.5194/amt-2017-395>, in review, 2018.](https://doi.org/10.5194/amt-2017-395)
- Giannakaki, E., Balis, D. S., Amiridis, V., and Zerefos, C.: Optical properties of different aerosol types: seven years of combined Raman-elastic backscatter lidar measurements in Thessaloniki, Greece, *Atmospheric Measurement Techniques*, 3, 569–578, <https://doi.org/10.5194/amt-3-569-2010>, 2010.
- [Giannakaki, E., Pfüller, A., Korhonen, K., Mielonen, T., Laakso, L., Vakkari, V., Baars, H., Engelmann, R., Beukes, J. P., Van Zyl, P. G., Josipovic, M., Tiitta, P., Chiloane, K., Piketh, S., Lihavainen, H., Lehtinen, K. E. J., and Komppula, M.: One year of Raman lidar observations of free-tropospheric aerosol layers over South Africa, *Atmospheric Chemistry and Physics*, 15, 5429–5442, doi:10.5194/acp-15-5429-2015, 2015.](https://doi.org/10.5194/acp-15-5429-2015)
- 35 [Giannakaki, E., Pfüller, A., Korhonen, K., Mielonen, T., Laakso, L., Vakkari, V., Baars, H., Engelmann, R., Beukes, J. P., Van Zyl, P. G., Josipovic, M., Tiitta, P., Chiloane, K., Piketh, S., Lihavainen, H., Lehtinen, K. E. J., and Komppula, M.: One year of Raman lidar observations of free-tropospheric aerosol layers over South Africa, *Atmospheric Chemistry and Physics*, 15, 5429–5442, doi:10.5194/acp-15-5429-2015, 2015.](https://doi.org/10.5194/acp-15-5429-2015)
- Giglio L., Schroeder, W., Justice, C.O. : The collection 6 MODIS active fire detection algorithm and fire products. *Remote Sensing of Environment*, 178, 31-41, 2016.

- Granados-Muñoz, M. J., Navas-Guzmán, F., Guerrero-Rascado, J. L., Bravo-Aranda, J. A., Biniotoglou, 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, *Atmos. Chem. Phys.*, **16**, 7043–7066, <https://doi.org/10.5194/acp-16-7043-2016>, 2016.
- 5 Groß, S., Tesche, M., Freudenthaler, V., Toledano, C., Wiegner, M., Ansmann, A., Althausen, D., and Seefeldner, M.: Characterization of Saharan dust, Maritime 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:10.1111/j.1600-0889.2011.00556.x, 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, <http://www.atmos-chem-phys.net/13/2487/2013/>, 2013.
- 10 Hamill, P., et al.: An AERONET-based aerosol classification using the Mahalanobis distance, *Atmos. Env.*, **140**, 213–233, doi:10.1016/j.atmosenv.2016.06.002, 2016.
- Hess, M., Koepke, P., and Schult, I.: Optical properties of aerosols and clouds: The software package OPAC, *B. Am. Meteorol. Soc.*, **79**, 831–844, [https://doi.org/10.1175/1520-0477\(1998\)0792.0.CO;2](https://doi.org/10.1175/1520-0477(1998)0792.0.CO;2), 1998.
- 15 Illingworth, A. J., Barker, H. W., Beljaars, A., Ceccaldi, M., Chepfer, H., Cole, J., Delanoë, J., Domenech, C., Donovan, D. P., Fukuda, S., Hirakata, M., Hogan, R. J., Huenerbein, A., Kollias, P., Kubota, T., Nakajima, T., Nakajima, T. Y., Nishizawa, T., Ohno, Y., Okamoto, H., Oki, R., Sato, K., Satoh, M., Shephard, M., Wandinger, U., Wehr, T., and Van Zadelhoff, G.-J.: THE EARTHCARE SATELLITE: The next step forward in global measurements of clouds, aerosols, precipitation and radiation, *Bulletin of the American Meteorological Society (BAMS)*, p. (in press), doi:10.1175/BAMS-D-12-00227.1, 2015.
- 20 Mahalanobis, P. C.: On the generalized distance in statistics. *Proceedings of the National Institute of Science of India*, **12**, 49–55, 1936.
- Mamouri, R. E. and Ansmann, A.: Fine and coarse dust separation with polarization lidar, *Atmos. Meas. Tech.*, **7**, 3717–3735, <https://doi.org/10.5194/amt-7-3717-2014>, 2014.
- Marinou, E., Amiridis, V., Biniotoglou, I., Tsikerdekis, A., Solomos, S., Proestakis, E., Konsta, D., Papagiannopoulos, N., Tsekeri, A., Vlastou, G., Zanis, P., Balis, D., Wandinger, U., and Ansmann, A.: Three-dimensional evolution of Saharan dust transport towards Europe based on a 9-year EARLINET-optimized CALIPSO dataset, *Atmospheric Chemistry and Physics*, **17**, 5893–5919, <https://doi.org/10.5194/acp-17-5893-2017>, 2017.
- 25 Matthias V., and Bosenberg J.: Aerosol climatology for the planetary boundary layer derived from regular lidar measurements, *Atmospheric Research*, Volume 63, Issues 3–4, Pages 221–245, ISSN 0169-8095, [https://doi.org/10.1016/S0169-8095\(02\)00043-1](https://doi.org/10.1016/S0169-8095(02)00043-1), 2002.
- 30 Mattis, I., Ansmann, A., Mueller, D., Wandinger, U. & Althausen, D.: Multiyear aerosol observations with dual-wavelength Raman lidar in the framework of EARLINET, *Journal of Geophysical Research: Atmospheres*, vol 109, no. D13, D13203. DOI: 10.1029/2004jd004600, 2004.
- 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, <https://doi.org/10.5194/acp-12-2229-2012>, 2012.
- 35 Mona, L., Papagiannopoulos, N., Basart, S., Baldasano, J., Biniotoglou, I., Cornacchia, C., and Pappalardo, G.: EARLINET dust observations vs. BSC-DREAM8b modeled profiles: 12-year-long systematic comparison at Potenza, Italy, *Atmos. Chem. Phys.*, **14**, 8781–8793, <https://doi.org/10.5194/acp-14-8781-2014>, 2014.

- Mona L., Kahn R., Papagiannopoulos N., Pappalardo G., Holzer-Popp T.: Aerosol typing—key information from aerosol satellite measurements, European Aerosol Conference, Milan (Italy) September 6-11, https://geko.promeeeting.it/abstract/DEF/Poster/2AAP_P094.pdf, 2015.
- 5 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: Micro-physical particle characterization, *J. Geophys. Res.*, **110**, D17201, 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, *Journal Of Geophysical Research*, Vol. 112, D16202, doi:10.1029/2006JD008292, 2007
- 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. Atmospheres*, **118**, 2956–2965, <https://doi.org/10.1002/jgrd.50324>, 2013.
- 10 Nicolae D., Vasilescu J., Talianu C. and Dandocsi A.: Independent Retrieval of Aerosol Type From Lidar, *EPJ Web of Conferences* **119**, 18002, <https://doi.org/10.1051/epjconf/201611918002>, 2016.
- Nicolae D., Vasilescu J., Talianu C., Biniotoglou I., Nicolae V., Andrei S., and Antonescu B., A Neural Network Aerosol Typing Algorithm Based on Lidar Data, *Atmos. Chem. Phys. Discuss.*, <https://doi.org/10.5194/acp-2018-492>, 2018
- 15 Nicolae D., Talianu C., Vasilescu J., Nicolae V., Stachlewska I.S.: Strengths and limitations of the NATALI code for aerosol typing from multiwavelength Raman lidar observations, *EPJ Web of Conferences* **176**, 05005, <https://doi.org/10.1051/epjconf/201817605005>, 2018
- Omar, A. H., Winker, D. M., Kittaka, C., Vaughan, M. A., Liu, Z., Hu, Y., Trepte, C. R., Rogers, R. R., Ferrare, R. A., Lee, K.-P., Kuehn, R. E., and Hostetler, C. A. : The CALIPSO automated aerosol classification and lidar ratio selection algorithm, *Journal of Atmospheric and Oceanic Technology*, **26**, 1994,doi:10.1175/2009JTECH1231.1, <http://ams.allenpress.com/perlserv/?request=get-abstract&.doi=10.2009>.
- 20 [Papagiannopoulos, N., Mona, L., Amodeo, A., D'Amico, G., GumClaramunt, P., Pappalardo, G., Alados-Arboledas, L., Guerrero-Rascado, J. L., Amiridis, V., Kokkalis, P., Apituley, A., Baars, H., Schwarz, A., Wandinger, U., Biniotoglou, I., Nicolae, D., Bortoli, D., Comerón, A., Rodríguez-Gómez, A., Sicard, M., Papayannis, A., and Wiegner, M.: An automatic observation-based aerosol typing method for EARLINET, *Atmos. Chem. Phys.*, **18**, 15879-15901, <https://doi.org/10.5194/acp-18-15879-2018>, 2018.](https://doi.org/10.5194/acp-18-15879-2018)
- 25 Papagiannopoulos, N., Mona, L., Alados-Arboledas, L., Amiridis, V., Baars, H., Biniotoglou, 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.
- 30 Papayannis, A., et al.: Systematic lidar observations of Saharan dust over Europe in the frame of EARLINET (2000–2002), *J. Geophys. Res.*, **113**, D10204, doi: 10.1029/2007JD009028, 2008.
- [Papayannis A., Mamouri R-E., Vassilis A., Giannakaki E., Veselovskii I., Kokkalis P., Tsaknakis G., Balis D., Kristiansen N., Stohl A., Korenskiy M., Allakhverdiev K.: 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 - ATMOS ENVIRON.* **48**, 10.1016/j.atmosenv.2011.08.037, 2012.](https://doi.org/10.1016/j.atmosenv.2011.08.037)
- 35 Pappalardo, G., A. Amodeo, L. Mona, M. Pandolfi, N. Pergola, and V. Cuomo(2004), Raman lidar observations of aerosol emitted during the 2002 Etna eruption, *Geophys. Res. Lett.*, **31**, L05120, doi: 10.1029/2003GL019073-, 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, *Atmospheric Chemistry and Physics*, 13, 4429–4450, doi:10.5194/acp-13-4429-2013, 2013.
- Pappalardo G., Amodeo A., Apituley A., Comeron A., Freudenthaler V., Linne H., Ansmann A., Bosenberg 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, 2014.
- Pérez, C., Nickovic, S., Pejanovic, G., Baldasano, J. M., and Özsoy, E.: Interactive dust-radiation modeling: A step to improve weather forecasts, *J. Geophys. Res.-Atmos.*, 111, d16206, <https://doi.org/10.1029/2005JD006717>, 2006b.
- 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.
- Russell, P.B., et al.: 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.
- Sawamura P., Vernier J.P., Barnes J.E., Berkoff T.A., Welton E.J., Alados-Arboledas L., Navas-Guzmán F., Pappalardo G., Mona L., Madonna F., Lange D., Sicard M., Godin-Beekmann S., and Payen G., Wang Z., Hu S., Tripathi S.N., and 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 034013, 2012
- Schmeisser, L., Andrews, E., Ogren, J. A., Sheridan, P., Jefferson, A., Sharma, S., Kim, J.E., Sherman, J. P., Sorribas M., Kalapov, I., Arsov, T., Angelov, C., Mayol-Bracero, O.L., Labuschagne, C., Kim, S.W., Hoffer, A., Lin, N.H., Chia, H.P., Bergin, M., Sun, J., Liu, P., and Wu, H., Classifying aerosol type using in situ surface spectral aerosol optical properties, *Atmos. Chem. Phys.*, 17, 12097–12120, <https://doi.org/10.5194/acp-17-12097-2017>, 2017.
- [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, PhD 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, *Atmos. Chem. Phys.*, 12, 3115–3130, <https://doi.org/10.5194/acp-12-3115-2012>, 2012a.](#)
- Siomos, N., Balis, D. S., Voudouri, K. A., Giannakaki, E., Filioglou, M., Amiridis, V., Papayannis, A., and Fragkos, K.: Are EARLINET and AERONET climatologies consistent? The case of Thessaloniki, Greece, *Atmos. Chem. Phys.*, 18, 11885–11903, <https://doi.org/10.5194/acp-18-11885-2018>, ~~2018~~-2018a.
- [Siomos N., Fountoulakis I., Natsis A., Karagiozidis D., Drosoglou T., Voudouri K.A., Bais A. and Balis D.: Aerosol classification based on the synergy of Brewer spectrophotometer and sunphotometer measurements, 14th International Conference on Meteorology, Climatology and Atmospheric Physics, <http://comecap2018.gr/wp-content/uploads/2019/03/COMECAP-BOA-final-revised.pdf>, 2018b.](#)

[Siomos, N., Climatology and classification of the optical and microphysical properties of aerosols using ground - based remote sensing techniques \(Doctoral dissertation\). Retrieved from http://hdl.handle.net/10442/hedi/45104, 2018c.](http://hdl.handle.net/10442/hedi/45104)

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, <http://dx.doi.org/10.1175/BAMS-D-14-00110.1>, 2015

5 Rolph, G., Stein, A., and Stunder, B.,. Real-time Environmental Applications and Display sYstem: READY. Environmental Modelling & Software, 95, 210-228, 2017.

Tesche, M., A. Ansmann, D. Müller, D. Althausen, R. Engelmann, V. Freudenthaler, and S. Groß, Vertically resolved separation of dust and smoke over Cape Verde using multiwavelength Raman and polarization lidars during Saharan Mineral Dust Experiment 2008, J. Geophys. Res., 114, D13202, doi:10.1029/2009JD011862, 2009a.

10 Tesche, M., Ansmann, A., Müller, D., Althausen, D., Mattis, I., Heese, B., Freudenthaler, V., Wiegner, M., Esselborn, M., Pisani, G., and Knippertz, P.: Vertical profiling of Saharan dust with Raman lidars and airborne HSRL in southern Morocco during SAMUM, Tellus B, 61, 144–164, <https://doi.org/10.1111/j.1600-0889.2008.00390.x>, 2009b.

Tesche, M., Detlef, M., 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, 15 <https://doi.org/10.3402/tellusb.v63i4.16362>, 2011a.

Tesche, M., Groß, S., Ansmann, A., Müller, D., Althausen, D., Freudenthaler, V., and Esselborn, M.: Profiling of Saharan dust and ~~biomass-burning~~ biomass burning smoke with multiwavelength polarization Raman lidar at Cape Verde, Tellus B, 63, <http://www.tellusb.net/index.php/tellusb/article/view/16360>, 2011b.

Wandinger, U., Baars, Holger, Engelmann, R., Hünerbein, A., Horn, S., Kanitz, T., Donovan, D., van Zadelhoff, G.-J., Daou, D., Fischer, J., 20 von Bismarck, J., Filipitsch, F., Docter, N., Eisinger, M., Lajas, D., and Wehr, T.: HETEAC: The Aerosol Classification Model for Earth-CARE, EPJ Web of Conferences, 119, 01 004, doi:10.1051/epjconf/201611901004, <http://dx.doi.org/10.1051/epjconf/201611901004>, 2016.

[Yorks J.E., Palm S.P., McGill, M.J., Hlavka D.L., Hart, W.D., Selmer P.A., and Nowotnick E.: CATS Algorithm Theoretical Basis Document, Level 1 and Level 2 Data Products, https://cats.gsfc.nasa.gov/media/docs/CATS_ATBD.pdf, 2016.](https://cats.gsfc.nasa.gov/media/docs/CATS_ATBD.pdf)

Table 1. Nomenclature of the aerosol types used in the High and Low resolution (without depolarization) mode for both automatic typing techniques and the correspondence with the ones used in this study.

NATALI		EARLINET Mahalanobis		This Study
<u>High Resolution</u>	<u>Low Resolution</u>	<u>High Resolution</u>	<u>Low Resolution</u>	<u>Low Resolution</u>
<u>Continental</u>	<u>Continental</u>	<u>CleanContinental</u>	<u>Continental</u>	<u>CleanContinental</u>
<u>ContinentaPolluted,</u> <u>ContinentalSmoke,</u> <u>Smoke,</u> <u>MixedSmoke/CC</u>	<u>Continental polluted,</u> <u>Smoke, Continental</u> <u>Smoke</u>	<u>Polluted Continental,</u> <u>Smoke</u>	<u>PollutedSmoke</u>	<u>PollutedSmoke</u>
<u>Dust, Volcanic,</u> <u>DustPolluted,</u> <u>MixedDust/CC,</u> <u>ContinentalDust,</u> <u>Mineral</u> <u>Mixtures/Volcanic</u>	<u>Dust</u>	<u>Dust, Volcanic, Mixed</u> <u>Dust, Polluted Dust</u>	<u>Dust, Polluted Dust</u>	<u>Dust</u>
<u>Marine,</u> <u>CoastalPolluted/CC,</u> <u>Coastal/CC</u>	<u>Marine</u>	<u>Mixed Marine</u>	<u>Mixed Marine</u>	<u>Maritime</u>

Table 2. Mean aerosol optical properties of the reference aerosol types used on the two automated algorithms.

NATALI	$AE_{550/350}$ $AE_{350/550}$	LR_{350}	LR_{550}	$CR_{350/550}$	$CR_{550/1000}$
Dust (D)	0.88–0.92	43–46	44–49	<u>1.51–1.55</u>	<u>1.1–1.14</u>
Continental-polluted <u>ContinentalPolluted</u> (CPolluted)	1.17–1.34	55–75	62–74	<u>1.34–2.29</u>	<u>1.33–1.65</u>
Smoke (S)	1.15–1.31	56–72	81–92	<u>1.90–2.59</u>	<u>1.52–1.61</u>
Clean -Continental (CC)	1.17–1.29	43–54	52–53	<u>1.56–2.07</u>	<u>1.37–1.85</u>
Maritime (MMM)	-0.26–0.21	13–32	19–25	<u>0.77–1.35</u>	<u>0.7–2.91</u>
EARLINET Mahalanobis distance-based typing algorithm	$AE_{355/1064}$ $BAE_{355/1064}$	LR_{532}	LR_{532}/LR_{355}		
Dust (D)	0.2 <u>0.01–0.8–1.0</u>	44 <u>1000</u>	41 <u>1000</u>	62–63	0.86 <u>0.77–1.34–1.04</u>
PollSmoke <u>PollutedSmoke</u> (PS)	1.1 <u>1.0–1.5–1.9</u>	52 <u>1000</u>	50 <u>1000</u>	82–97	0.68 <u>0.87–1.12–1.6</u>
Clean-Continental <u>CleanContinental</u> (CC)	0.8 <u>0.7–1.4–1.6</u>	35 <u>1000</u>	32 <u>1000</u>	47–44	0.60 <u>0.62–1.0–1.04</u>
Maritime (MMM)	0.7 <u>0.5–1.1–1.2</u>		16–32		0.8 <u>0.81–1.0–0.9</u>

2013-05-20 Thessaloniki

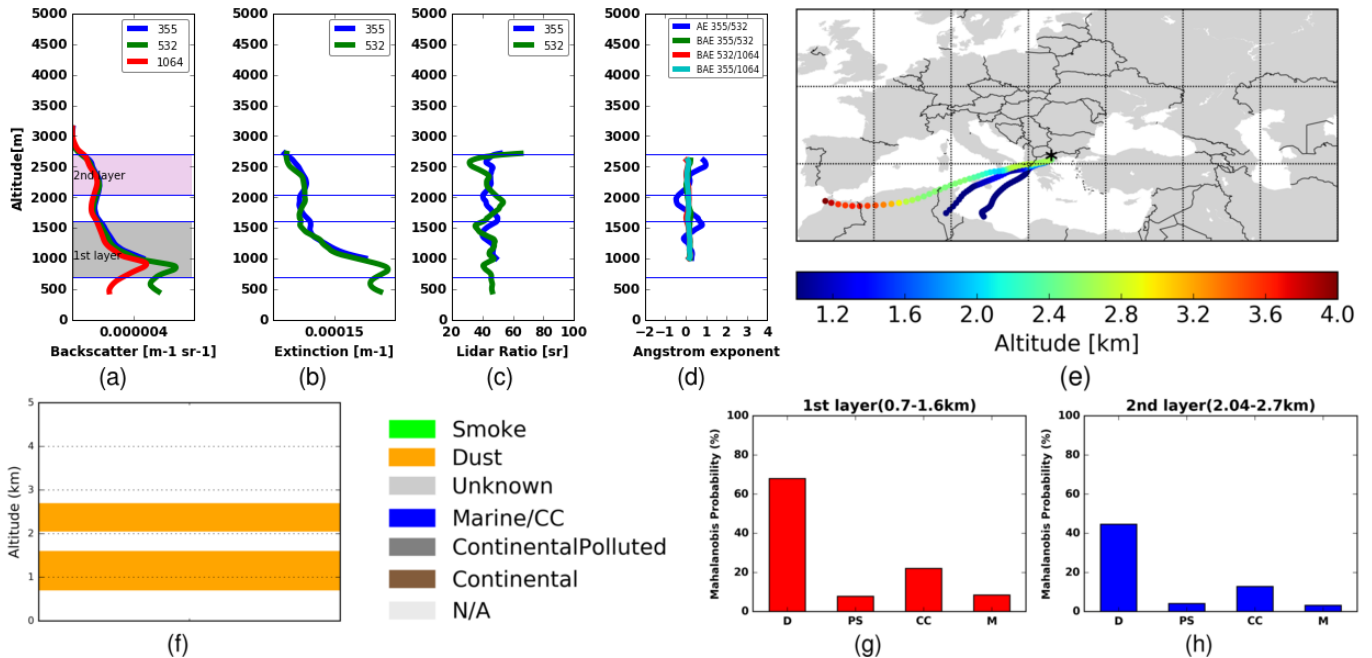


Figure 1. Vertical profiles of aerosol optical properties for Thessaloniki on 20th May 2013. (a) Backscatter coefficient at 355nm, 532nm and 1064nm, (b) extinction coefficient at 355 nm and 532nm, (c) Lidar Ratio, (d) Ångström exponent, (e) ~~seven-days~~ 4-days HYSPLIT backward trajectories arriving at Thessaloniki at each detected layer height, (f) aerosol main layers typed by NATALI and (g,h) EARLINET Mahalanobis distance-based typing algorithm probability for each detected layer (different from the normalized one).

2013-09-02 Thessaloniki

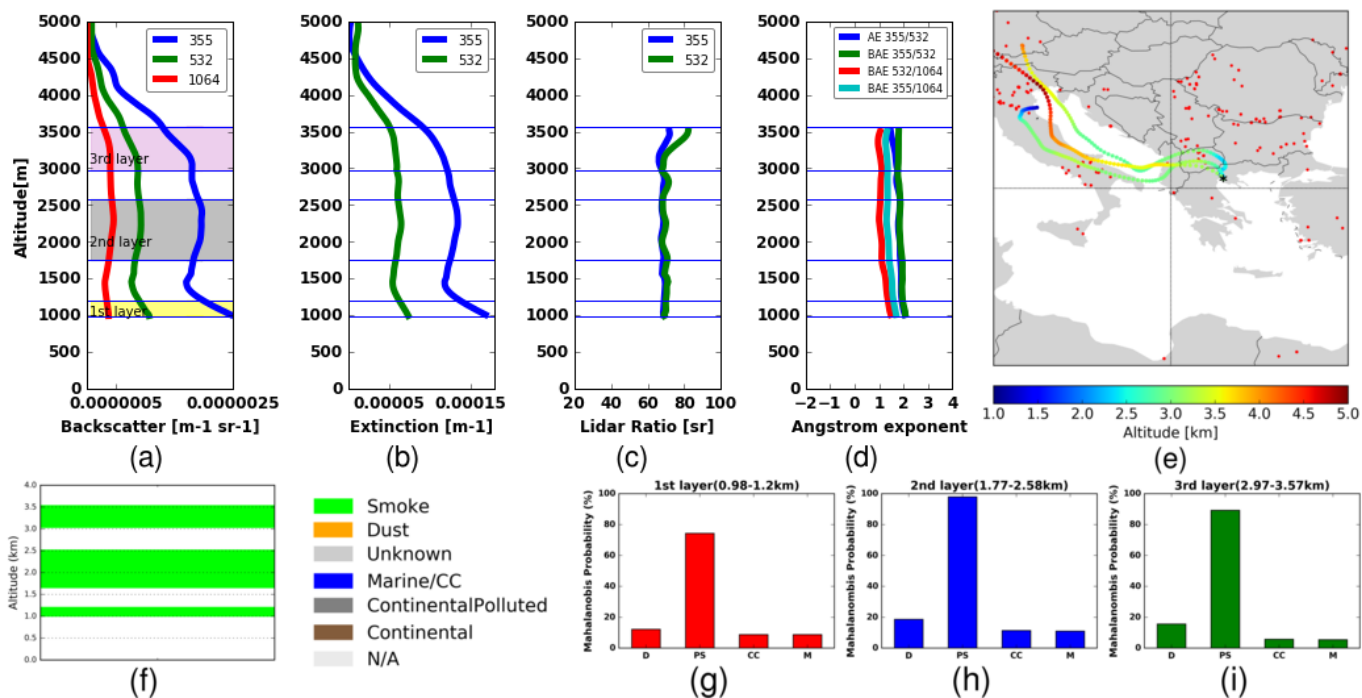


Figure 2. Vertical profiles of aerosol optical properties for Thessaloniki on 2nd of September 2013. (a) Backscatter coefficient profiles at 355nm, 532nm and 1064nm, (b) extinction coefficient profiles at 387nm and 607nm, (c) Lidar Ratio, (d) Ångström exponent, (e) [seven-days 3-days](#) backward trajectories arriving at Thessaloniki at each detected layer height, (f) aerosol main layers typed by NATALI, and (g,h,i) EARLINET Mahalanobis distance-based typing algorithm probability for each detected layer ([different from the normalized one](#)).

2014-08-21 Thessaloniki

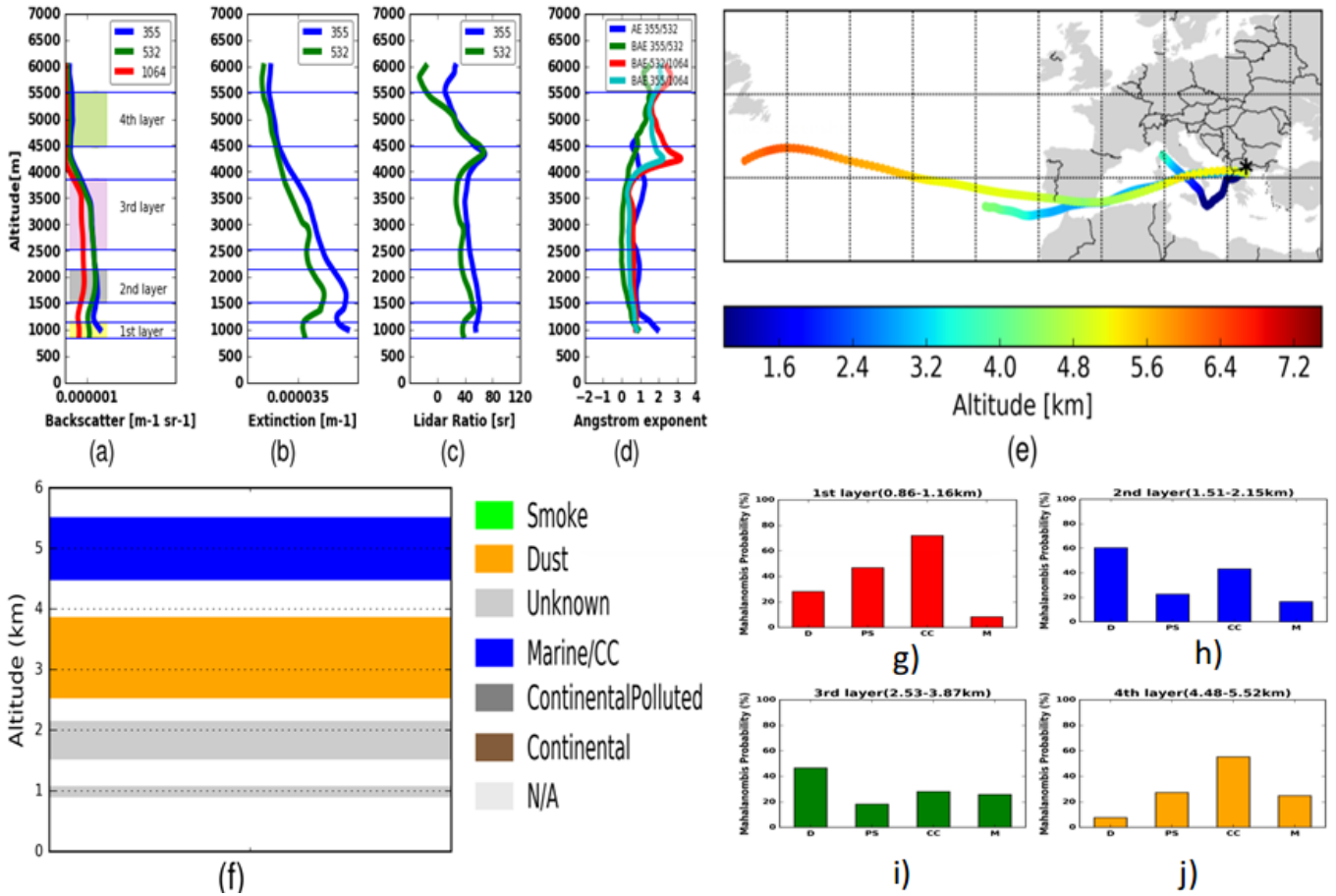
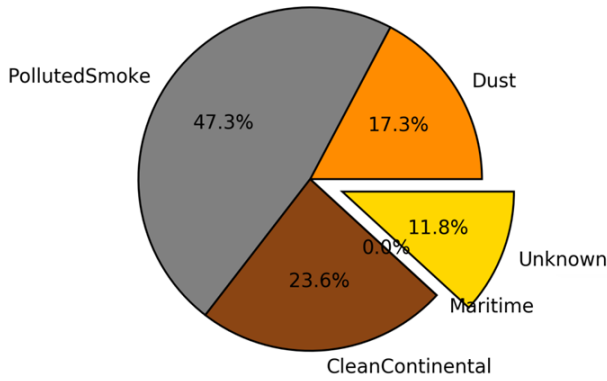


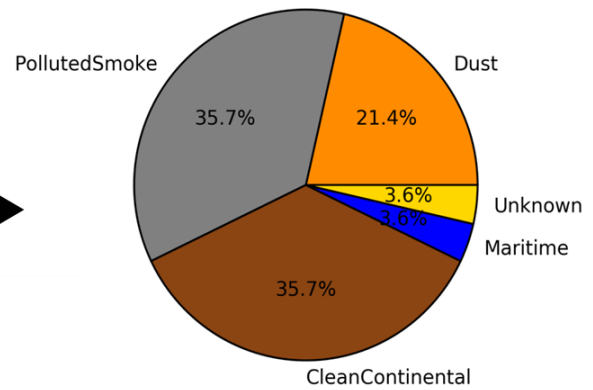
Figure 3. Bar-plots showing the number Vertical profiles of detected layers per aerosol class from both algorithms optical properties for Thessaloniki station on 21st of August 2014. (a) Backscatter coefficient profiles at 355nm, 532nm and 1064nm, (b) extinction coefficient profiles at 387nm and 607nm, (c) Lidar Ratio, (d) Ångström exponent, (e) 5-days backward trajectories arriving at Thessaloniki at each detected layer height, (f) aerosol main layers typed by NATALI, and (g,h,i,j) EARLINET Mahalanobis distance-based typing algorithm probability for the period under study each detected layer (2012–2015 different from the normalized one).

Aerosol Typing by Mahalanobis



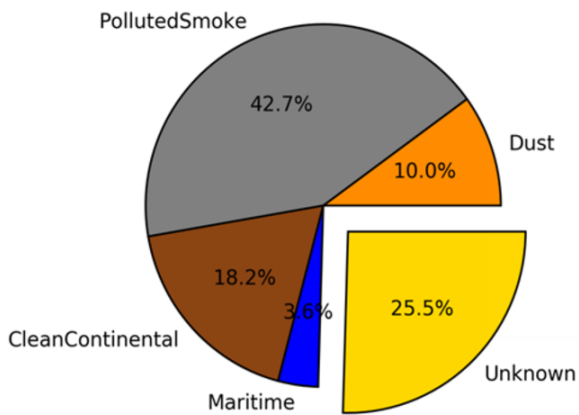
a)

U/N layers from NATALI Typed by Mahalanobis



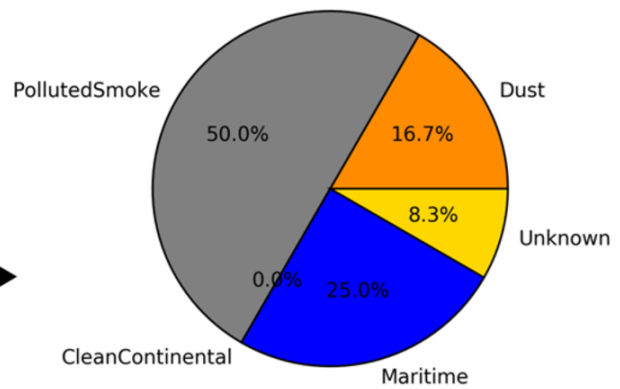
b)

Aerosol Typing by NATALI



c)

U/N layers from Mahalanobis Typed by NATALI



d)

Figure 4. Prediction accuracy for each aerosol class derived by the training Percentages of the EARLINET Mahalanobis distance-based typing algorithm to NATALI modeled data set typed and untyped layers for both automated algorithms.

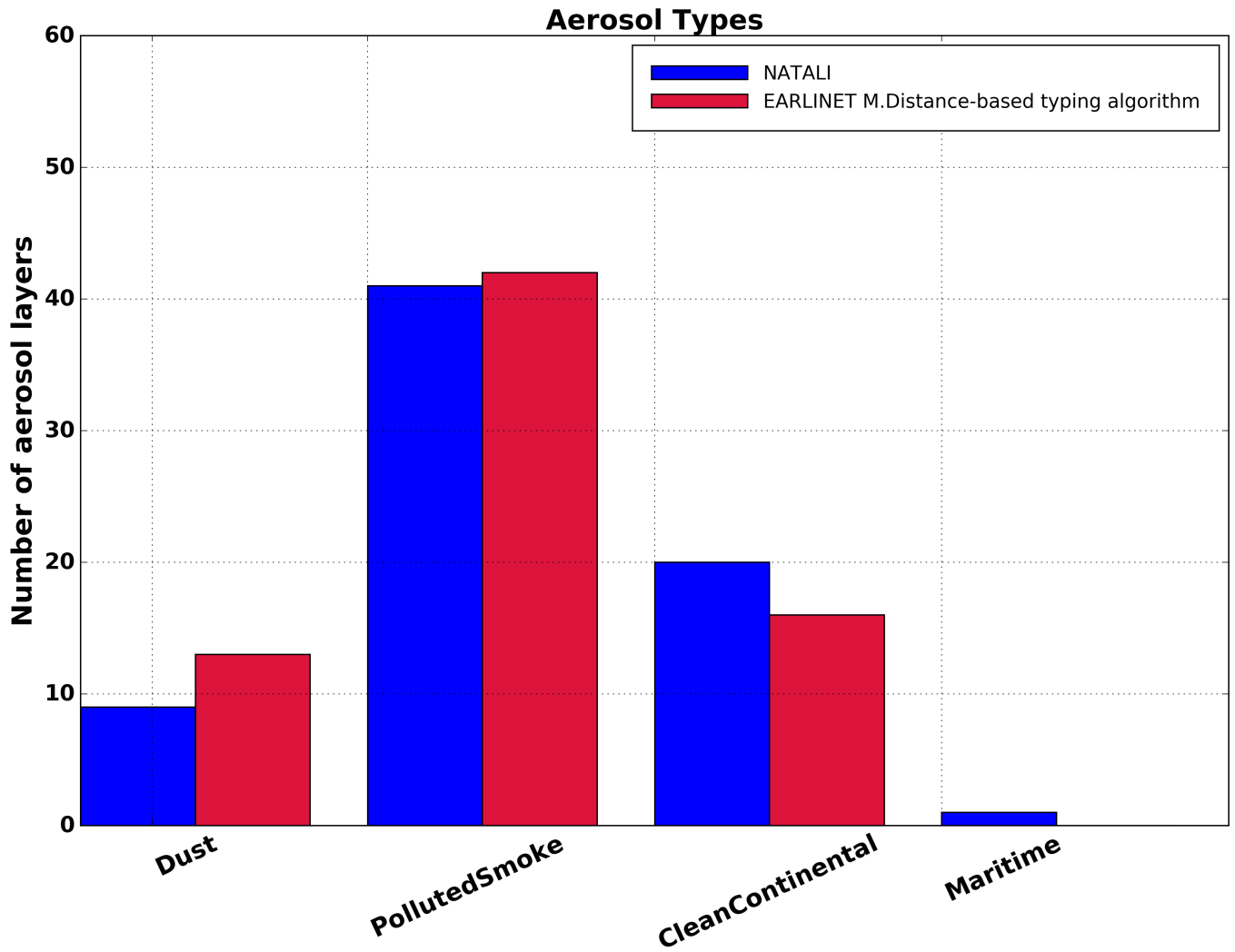


Figure 5. Prediction accuracy for each aerosol class derived by the two algorithms for Thessaloniki station for, based on the period-under-study 71 cases that both algorithms provided results.

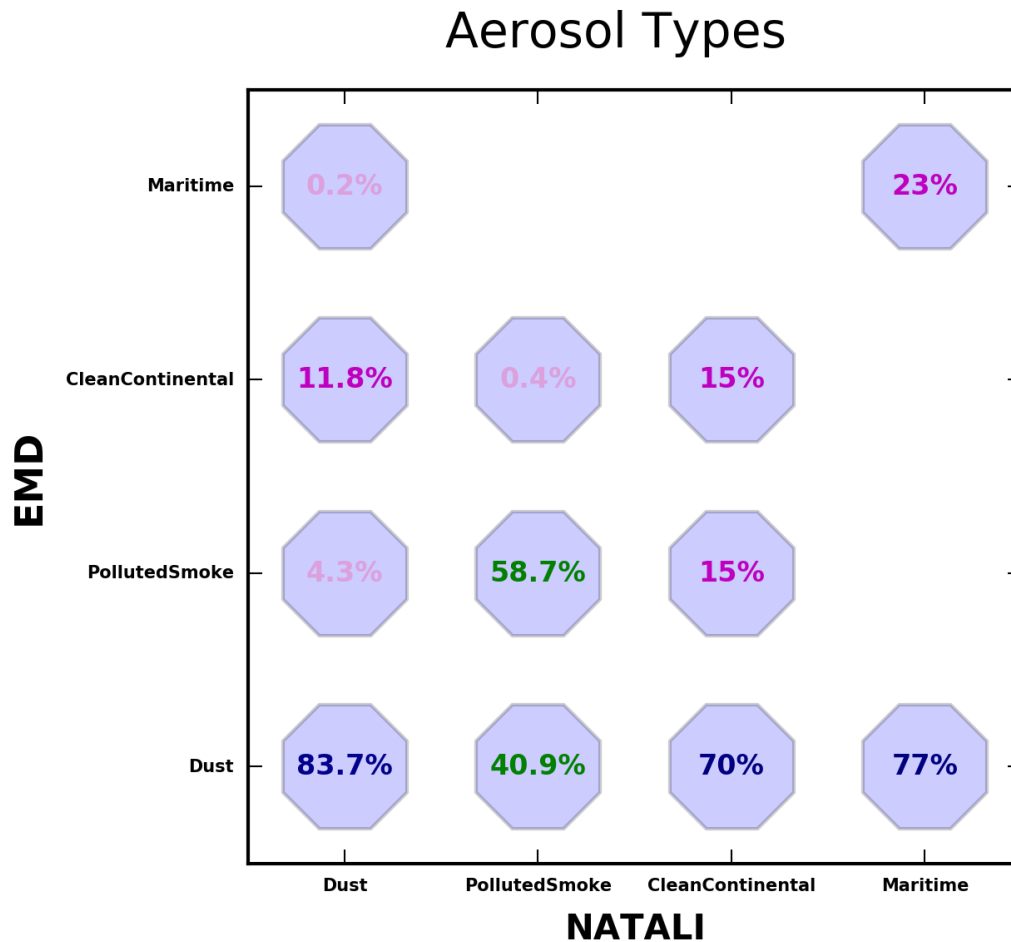


Figure 6. Typing score of each aerosol class derived by the EARLINET Mahalanobis distance-based typing algorithm (EMD), when applied to NATALI modeled data set (total of 259407 cases). The denominator in each column is the number of NATALI samples for a given type, and the numerator is the number of the EARLINET Mahalanobis distance-based classifications for the same sample.

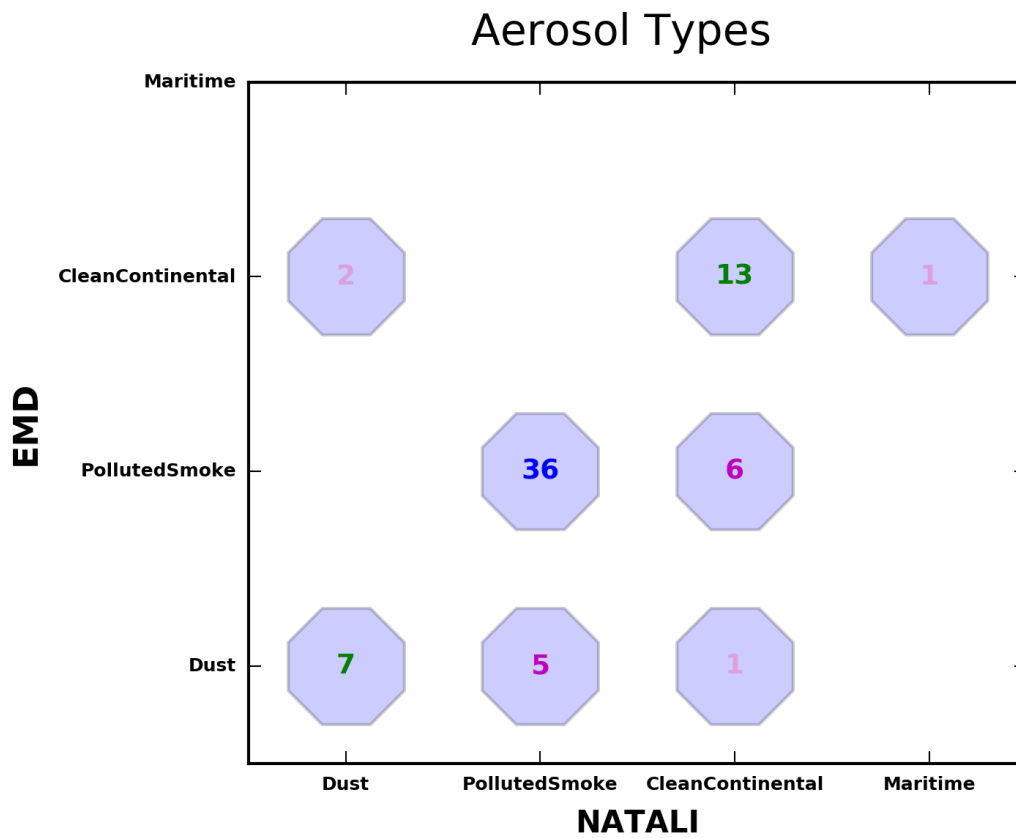


Figure 7. Number of cases per aerosol type derived by the two algorithms (EARLINET Mahalanobis distance-based typing algorithm - EMD and NATALI) for the period 2012-1015 (71 cases).