Response to Referee 1

We would like to thank the reviewer for his/her fruitful comments that helped to improve the manuscript.

The paper of Siomos et al. presents an interesting example of aerosol model evaluation based on remote sensing observations. The manuscripts highlights the potential and pitfalls for such a comparison, therefore it could be of interest for the wider atmospheric community. The manuscript is worth publishing after addressing several comments listed below.

General comments

1) A main issue with the presented analysis is that the authors compare fine and coarse particles defined in two fundamentally different ways. According to the text, the model's fine mode is defined as particles with aerodynamic diameter less than 2.5um, while LIRIC's fine mode is defined as particles with (optical) diameter less than ~0.4 – 1.2 um. Before this study is published, the authors should thoroughly discuss this issue and justify why their comparison gives any meaningful results.

The reviewer is right. In the current analysis the PM2.5 particles should include all the fine particles and a small part of the coarse particles that is variable depending on the case. After analyzing the size distribution of all the cases, we found that this fraction of the coarse mode ranges from 5-25%. It is possible to convert the fine and coarse modes of LIRIC to PM2.5 and PM2.5-10 particles using this fraction. In the LIRIC inversion, the normalized size distribution of each mode is derived from the columnar size distribution for each height bin, resulting in constant extinction and backscattering efficiencies per aerosol mode. Taking that into account, the fraction of the sunphotometer's coarse mode that belongs in the PM2.5 region is independent of the height. Thus, it is possible to use this fraction in order to convert the LIRIC fine and coarse profiles to PM2.5 and PM2.5-10 profiles that are consistent with CAMx, by subtracting (for each individual case) the PM2.5 coarse fraction from each LIRIC coarse profile and adding it to the respective LIRIC fine profile. This affects the LIRIC fine and coarse concentration and integrated mass values as well as the fine center of mass values. The manuscript has been updated accordingly. The "fine" and "coarse" terminology has been replaced by "PM2.5" and "PM2.5-10" where it was necessary. The tables 3, 4, 6 and figures 3, 4, 5, 6, 7 are modified. The discussion and is also modified accordingly. The following paragraphs have been added to the text to describe this methodology.

The following text was added at the end of Section 3.1.1: “Another hindrance in the analysis is that the fine and coarse mode of LIRIC are not directly comparable with the PM2.5 and PM2.5-10 modes of CAMx. The PM2.5 particles should include all the fine particles and a small part of the coarse particles that changes depending on the case. Additionally, the size distribution of the sunphotometer usually surpasses the PM10 diameter limit. Fortunately, it is possible to convert the fine and coarse modes of LIRIC to PM2.5 and PM2.5-10 particles. In the LIRIC inversion, the normalized volume size distribution of each mode is derived by separating the columnar size distribution of the sunphotometer in the two modes. The normalized distribution of each mode remains constant with height. Taking that into account, the fractions of the sunphotometer’s coarse mode that belong in the PM2.5 region and the region outside the PM10 particles can be calculated from the sunphotometer's volume size distribution. Then, the fine and coarse concentration profiles of each LIRIC case can be converted to PM2.5 and PM2.5-10 profiles using the equations 2 and 3. ”

The new equations are presented below:

$$c_{PM_{2.5}}(z) = c_{fine}(z) + c_{coarse}(z) \cdot \frac{\int_{r_{f-c}}^{r_{PM_{2.5}}} \frac{dV}{dr} \cdot dr}{\int_{r_{f-c}}^{r_c} \frac{dV}{dr} \cdot dr}$$

$$c_{PM_{2.5-10}}(z) = c_{coarse}(z) - c_{coarse}(z) \cdot \frac{\int_{r_{f-c}}^{r_{PM_{2.5}}} \frac{dV}{dr} \cdot dr + \int_{r_{PM_{10}}}^{r_c} \frac{dV}{dr} \cdot dr}{\int_{r_{f-c}}^{r_c} \frac{dV}{dr} \cdot dr})$$

Where c_{fine} , c_{coarse} , $c_{PM2.5}$, $c_{PM2.5-10}$ are the concentration profiles of LIRIC before and after the conversion and dV/dr is the aerosol volume size distribution of the sunphotometer as a function of the aerosol radius. The radii r_c , r_{f-c} , $r_{PM2.5}$, r_{PM10} are in μm units and correspond to the upper limit of the sunphotometer size distribution, the separator radius between the fine and the coarse mode of the sunphotometer and the PM2.5 and the PM10 separator radii respectively.

Concerning the optical versus the aerodynamic diameter, it is possible to convert from one type to the other (C.-H. Chien et al. / Journal of Aerosol Science 101 (2016) 77–85). Using their formula for the NaCl particles, an aerodynamic diameter of 2.5um

corresponds to an optical diameter of approximately 2.0um. However, we aren't going to perform this conversion for the following reasons.
The aerosol concentration in CAMx depends on the emissions within the domain and the boundary conditions. In both cases, the species concentration is imported from external models (TNO emissions, ECMWF emissions, NEMO). In general, the aerosol concentration in models is based on satellite and ground based measurements. Taking this into account, it is difficult to characterize the aerosol diameter of a model as exclusively optical or aerodynamic. Even the particles that are produced from chemical reactions inside CAMx (i.e. the secondary organic compounds) do not carry the information of a detailed size distribution. They are just flagged as PM2.5 or PM2.5-10. As a result, we removed the word "aerodynamic" from the manuscript since it is misleading.

2) Desert dust is included in the model only as a boundary conditions and this explains, according the the authors, the poor performance of the model in forecasting coarse aerosol concentration. However most desert dust is produced outside the model's domain. Given appropriate boundary conditions, CAMx should transport the dust in its domain and produce good prediction of dust concentration. Do the author's imply that the MACC models provide bad boundary conditions or does CAMx do a poor job transporting the dust within its domain?

Following the reviewer's suggestion we examined maps of CAMx for selective cases that were affected by desert dust and it seems that, for some of them there are issues in the transportation of CAMx PM2.5-10 from the boundaries to long distances. Taking into account the number of dust cases in this study, it is difficult to draw a firm conclusion on the prevalent source of bias. This can be examined in a future study. Consequently, the text in section 4.1, 4.2 and in the conclusions is modified so that both the lack of dust emissions in the domain (other than the boundary conditions), and the model's transportation of dust are presented as potential sources of bias in the dust concentration.

During the analysis we detected a bug in our algorithms where the "soil PM2.5" and the "soil PM2.5-10" components were in some cases identified as the "other PM2.5" and the "other PM2.5-10" components respectively and vice versa. This is now corrected and the relevant discussion in both sections 4.1 and 4.2 is modified accordingly. The new figures 4c, 4d and 7 (left) as well as tables 5 and 6 are modified accordingly.

In addition, after reinspecting the data processing algorithms, we noticed that for the center of mass calculation the vertical resolution of the profiles was considered constant which is not the case when the model's eta levels are considered. We recalculated the center of mass with variable vertical resolution and tables 3 and 5 and figures 4 and 7a have been modified accordingly.

3) Section 2.5 should define the uncertainties of the LIRIC algorithm. Several references to evaluation studies are given in the last paragraph, but the authors should briefly present the outcome of these studies, at least to the extent that are relevant for the discussion of their results.

The text was modified to: "The effects of multiple user defined uncertainties, such as the upper and lower limit heights of the profile and the algorithm's regularization parameters, on the final result has been studied by Granados-Muñoz et al. (2014) and Filioglou et al. (2017) for selective case studies in Granada and Thessaloniki respectively. They agree that the parameter that produces the biggest uncertainties is the lower limit height of the profile. Furthermore, the LIRIC retrievals have already been evaluated for volcanic and desert dust particles by Wagner et al. (2013) showing that the inversion can be accurate for two quite different types of aerosol. The aerosol extinction products of LIRIC has also been compared against the respective products from the Generalized Aerosol Retrieval from Radiometer and Lidar Combined data (GARRLiC) algorithm and against the retrievals from raman lidar measurements (Bovchaliuk et al., 2016). Finally, LIRIC has also been validated against in-situ aircraft measurements (e.g., Granados-Muñoz et al., 2016a; Kokkalis et al., 2017). Granados-Muñoz et al. (2016a) compared the LIRIC retrievals with airborn in-situ measurements and found a promising agreement with the differences between the two staying within the expected uncertainties. Kokkalis et al. (2017) analyzed a biomass burning case. Their comparison between the LIRIC retrievals and the aircraft measurements resulted in a good performance of the algorithm for the fine particles. As a result it can be used as an independent reliable tool for the validation of CAMx."

4) The author's definition of PBL is not consistent with the description of the LIRIC algorithm. The authors claim that they search for PBL's top height between "400m and 2.5km". However, LIRIC's lower boundary is set to 600m.

Indeed this is not clear in the text. The overlap correction is applied normally but it is still necessary to limit the profiles since the overlap function can't be trusted down to the ground. For the comparison between LIRIC and CAMx we chose to limit the profiles in a region where the overlap function is above 0.9 (600m) instead of 0.7 that we typically use in the lidar data processing in order to reduce the uncertainty of the overlap correction. The 600m limit also apply to the PBL height retrieval. The 400m is a typo and it will be corrected.

The text was modified according to the reviewer's suggestions: "Identification criteria are necessary for the selection of the PBL height. The top of the layer between 600m and 2.5km with the minimum value in the transformed signal is chosen as the boundary layer height."

In addition LIRIC is "demanding a certain degree of vertical smoothness in the final product", possibly masking the true PBL top. The authors should address these discrepancies and provide estimated of the resulting uncertainties.

By comparing the Klett lidar backscatter profiles from our operational algorithms and the LIRIC backscatter profiles using the concentration and the backscattering efficiencies from LIRIC (see equation) we have seen that the vertical structure is similar, especially for strong layers such as the boundary layer.

They should also compare the PBL values derived from LIRIC with the PBL values assumed in the corresponding model profiles.

Since the vertical resolution is quite lower in CAMx than in LIRIC (eta levels against a constant vertical resolution of 15m) it would be pointless to apply the WCT or similar techniques that take advantage of the aerosol vertical distribution to the profiles of CAMx.

Technical corrections

Page 1
Line 1: missing parenthesis " with extensions (CAMx)." This applies also to page 2, line 3.

The text was changed according to the reviewer's suggestion.

Line 2: "updated version of the former". This is awkward wording.

The text was rephrased to: "the Dust Regional Atmospheric Model (BSC-DREAM8b) were deployed"

Page 2

Line 13: “For example Mona et al. (2014) compare [..] the dust extinction”. Delete “be- tween”.

The text was changed according to the reviewer’s suggestion.

Line 31: “(ENVIRON, 2010)”

The text was changed according to the reviewer’s suggestion.

Page 3

Line 27: “Schneider et al. (2000)”. The citation seems misplaced and poorly formated.

The text was changed to: “(EARLINET) (Schneider et al., 2000; Papalardo et al., 2014)”

Page 7

Line 20: “In the current dataset the full overlap height was calculated at 900m. The lower boundary is set to 600m where the overlap function is still above 90%.”. Provide more information about these calculations.

See the relative comment response in Reviewer 3.

Page 7, Line 20: The text has been modified to: “A lower height boundary has to be determined due to the overlap function of the lidar system. Operationally, we apply the method of Wandinger et al. 2002 for the calculation of the overlap function and the full overlap height. In the current dataset the full overlap height was calculated at 900m. The correction however cannot be trusted down to the ground (Wandinger et al. 2002). In this study, we apply the correction down to 600m where the overlap function is still above 90% and use this height as the lower boundary of the LIRIC inversion. Below this height the lidar signals are considered constant during the LIRIC inversion. The concentration retrievals are also kept constant below 600m.”

Page 10

Line 3: How are Q factors calculated?

The text was rephrased to: "where Q ext is the extinction efficiency and Q bsc is the backscattering efficiency calculated by LIRIC."

Many citations are badly formatted and need to be corrected.

The urls in the citations were removed. Some empty fields were also cleared.

Table 2, caption: "The a and c symbols". "a" should be "z".

The text was changed according to the reviewer's suggestion.

Fig. 2: What is the meaning of black dots in the HYSPLIT plots?

The following sentence was added in the caption of Figure 2: "The big black dots in a and b indicate 24h intervals"

Response to Referee 2

We would like to thank the reviewer for his/her fruitful comments that helped to improve the manuscript.

General comments: An evaluation study is presented to assess the capability of the air quality model CAMx to describe the aerosol conditions over Thessaloniki. The model simulations are compared to combined sun photometer and lidar observations. Backward trajectories and results of a sophisticated dust forecast model are used to attribute shortcomings to a poor representation of biomass burning and desert dust aerosol. In principle, I like the idea of using different tools, not only measurements, to evaluate the simulations of a specific model and track down shortcomings to suggest model improvements. The evaluation is properly done, although the focus on the comparison with LIRIC data from Thessaloniki only may be too one-sided.

Maybe other observations could be additionally included to underpin the findings.

Unfortunately there weren't any other LIRIC estimates available from lidar stations that are included in the modelling domain.

However, my essential criticism is that the CAMx model is evaluated regarding two aerosol types, which, by design, are not directly computed or only poorly represented. Biomass burning emissions are highly variable in time and space. The actual pollution will largely depend on specific events. Of course, it is not to be expected that the TNO emission database from 2007 in detail is representative for the fire emissions in 2013 – 2015. The same holds for Saharan dust that is not online computed based on modelled winds but input as boundary condition. This must be considered when evaluating the model results, and the conclusions have to be revised in this regard. How exactly is the CAMx model suggested to be improved with this in mind, and based on the evaluation results?

Our aim was not to evaluate CAMx for its performance regarding smoke and desert dust. At a first step we tried to use all available measurements for the period under study in order to investigate whether there is a good agreement between the model and the LIRIC estimates. From the analysis we concluded that the agreement is very good for fine mode aerosols excluding the smoke incidents and dust events, which means that most other sources (anthropogenic and natural) are reasonably represented in the model. Concerning the smoke we suggest that we cannot expect an agreement, since the emission inventory has not such on-line module. Concerning the

Saharan dust, indeed any desert dust in CAMx simulations results from the boundary conditions. Following the reviewer's suggestion we examined maps of CAMx for selective cases that were affected by desert dust and it seems that, for some of them there are issues in the transportation of CAMx PM2.5-10 from the boundaries to long distances. However, the small number of cases available for such an analysis does not allow to draw firm conclusions on this issue, especially to distinguish what is the main issue, the boundaries themselves or the transport. A relevant discussion is added in sections 4.1 and 4.2 and in the conclusions. The figures 3c and 3d and 6 on the left as well as tables 5 and 6 were also modified.

Specific comments:

1. Page 4, line 10: A plot showing the model domains would be very helpful, in particular, to show if relevant Saharan dust sources are included.

The domains of CAMx have been included in the manuscript (Figure 1). The figure numbering has been adjusted in the text.

The text has been modified according to the reviewer's suggestion: "The domains of CAMx are presented in figure 1."

2. Page 11, lines 5 – 8: Here and later in Section 4, the study period 2013 – 2015 should be mentioned in order to clearly separate example cases from the broader statistical analysis.

The text was modified according to the reviewers suggestion:

Page 11, line 5: "An ensemble of 24 measurements in the period 2013-2015."

Page 11, line 10: "In this section the simulated profiles of CAMx are compared against the observational profiles of LIRIC in the period 2013-2015."

3. Figures 1 and 3 – 7: Please indicate in each figure caption whether the results refer to a specific case or the entire period 2013 – 2015.

The text was modified according to the reviewers suggestion.

The phrase "the period 2013-2015" was added to all the aforementioned figures.

Response to Referee 3

We would like to thank the reviewer for his/her fruitful comments that helped to improve the manuscript.

The study presents an evaluation of CAMx model against LIRIC output profiles retrieved above the city of Thessaloniki. It is an interesting study with valuable results for the scientific community. However, the authors need to address some issues before publication. As it is currently presented, the idea of the validation is sometimes lost along the manuscript and the paper becomes a little too descriptive. The manuscript would benefit from a more in-depth discussion regarding the validation and more discussion including uncertainties is definitely needed. A review of the writing, which is sometimes confusing, and a possible shortening in length would also be useful to improve the manuscript. Find some more detailed comments below:

<u>Page 1</u>
Line 7: A fractional bias of 24.8% does not seem “close”. I suggest you use the absolute value here instead of percentage.

The text has been modified to: “mean bias of 0.57 km.”

<u>Page 2</u>
Line 3: Rephrase this sentence. As it is written, it looks like EMEP is a model instead of a programme.

The following text was removed: “European Monitoring and Evaluation Programme EMEP”

Lines 27-35: The identification of PM2.5 and PM10 particles with the fine mode and the coarse mode from LIRIC is not completely accurate. Please, rewrite.

See the relative comment response in Reviewer 1.

Page 2, Line 34: The text was modified to: “Instead of evaluating the performance of CAMx only for the PM10 particles, we separate the fine from the coarse particles by applying the LIRIC technique, then we convert the fine and coarse concentration profiles of LIRIC to PM2.5 and PM2.5-10 profiles and perform the validation for the PM2.5 and PM2.5-10 individually.”

Page 3

Lines 1-9: This information seems more appropriate for the methodology section than for the introduction.

In the introduction we briefly mention the tools used in our study which are described in more detail in the methodology section.

Line 25: “pre-processing”

The text has been modified according to the reviewer’s suggestion.

Line 27: Parenthesis are missing for the reference Schneider et al. (2000). Please, also add the more recent reference Pappalardo et al. (2014)

The text was changed to: “(EARLINET) (Schneider et al., 2000; Papalardo et al., 2014)”

Line 31: Was the sun photometer deployed at Thessaloniki just for this study?

The lines 31-32 have been rephrased to: “We used measurements from a CIMEL multiband sun-sky photometer which was installed in Thessaloniki in 2003 as part of the AERONET Global Network.”

Page 6

Line 21: Please, rewrite. It is not clear what you mean by “user defined uncertainties”. Does the study by Filioglou et al. (2016) take into the account the uncertainties in the input lidar and radiometer data or just the user defined input parameters? In that case, what is the estimated uncertainty of the output profiles? Include also here that LIRIC has been validated against in-situ aircraft measurements to emphasize that it can be used as an independent reliable tool for the validation of CAMx (see e.g. Granados-Munoz et al., 2016 and Kokkalis et al., 2017)

See the relative comment response to Reviewer 1.

The text was modified to: “The effects of multiple user defined uncertainties, such as the upper and lower limit heights of the profile and the algorithm’s regularization parameters, on the final result has been studied by Granados-Muñoz et al. (2014) and Filioglou et al. (2017) for selective case studies in Granada and Thessaloniki respectively. They agree that the parameter that produces the biggest uncertainties is the lower limit height of the profile. Furthermore, the LIRIC retrievals have already been evaluated for volcanic and desert dust particles by Wagner et al. (2013) showing that the inversion can be accurate for two quite different types of aerosol. The aerosol extinction products of LIRIC has also been compared against the respective products from the Generalized Aerosol Retrieval from Radiometer and Lidar Combined data (GARRLiC) algorithm and against the retrievals from raman lidar measurements (Bovchaliuk et al., 2016). Finally, LIRIC has also been validated against in-situ aircraft measurements (e.g., Granados-Muñoz et al., 2016a; Kokkalis et al., 2017). Granados-Muñoz et al. (2016a) compared the LIRIC retrievals with airborn in-situ measurements and found a promising agreement with the differences between the two staying within the expected uncertainties. Kokkalis et al. (2017) analyzed a biomass burning case. Their comparison between the LIRIC retrievals and the aircraft measurements resulted in a good performance of the algorithm for the fine particles. As a result it can be used as an independent reliable tool for the validation of CAMx."

line 29: What do you mean by characterization procedure of the lidar profiles?

The text has been changed to: "aerosol type identification"

Page 7
Line 20: How did you calculate the full overlap height? Add references here and/or provide more details.

From the method of Wandinger et al. 2002 both the overlap function and the full overlap height are calculated. In this study we have applied the overlap correction per case using a typical overlap function. The overlap correction, however, cannot be extended to the ground. Typically, we limit the profile at the height where the function is higher than 0.7. For the CAMx validation however we preferred to use 0.9 (600m) to be on the safe side. This not clear in the text and it will be added.

Furthermore, in the original analysis we kept the lidar signals constant below 600m during the LIRIC inversion but the concentration product of LIRIC can be slightly variable even below the lower limit. To be entirely consistent with the idea of constant products below 600m we decided to keep the concentration profiles constant below this lower limit. This slightly affects the figures 3,4,5,6 and the tables 3,4,5,6. The text was also modified in order to clarify this adjustment.

Page 7, Line 20: The text has been modified to: "A lower height boundary has to be determined due to the overlap function of the lidar system. Operationally, we apply the method of Wandinger et al. 2002 for the calculation of the overlap function and the full overlap height. In the current dataset the full overlap height was calculated at 900m. The correction however cannot be trusted down to the ground (Wandinger et al. 2002). In this study, we apply the correction down to 600m where the overlap function is still above 90% and use this height as the lower boundary of the LIRIC inversion. Below this height the lidar signals are considered constant during the LIRIC inversion. The concentration retrievals are also kept constant below 600m."

Line 26: Be more specific for the maximum height, what it is consider a significant quantity?

The text was rephrased to: "The upper boundary depends on the maximum height where aerosol exist in a significant quantity, that is, a region where the lidar signal from the aerosol backscattering can no longer be separated from the noise. This height can vary depending on the atmospheric conditions."

Line 27: Replace summing by adding

The text was modified according to the reviewer's suggestion.

Page 8
Line 5: Why are you using 1.5 and 2.6 g*cm-3? Why don't you use the known aerosol densities provided by CAMx for each case? That would lead to a more accurate comparison between LIRIC and CAMx.

The use of a CAMx derived density on the LIRIC profiles presupposes that the mixing ratio of each species is well predicted by CAMx. Otherwise the reference data of LIRIC would be affected by uncertainties originating from CAMx. Thus we preferred to use constant conversion values that are commonly used in the literature. This is also the way of Binietoglou et al. 2015. A more direct comparison would be to convert the CAMx profiles to ppbv. This conversion, however, was rejected after testing it because we wanted to avoid confusion by using a unit that is not adopted by the modeler's community like ppbv and the results were also pretty similar.

Lines 8-14: Since CAMx lacks of biomass burning aerosol emissions and does not consider desert dust emissions directly, I understand that the fires and dust categories are only used to evaluate the impact that this cases have on the model performance. However, for the evaluation purpose it would make more sense to me to include a category excluding biomass burning and dust cases. That way you would be comparing apples to apples.

The main reason that we didn't isolate the cases that aren't biomass burning and dust is that our dataset is limited. By removing the 6 dust cases from the "non fires" group, the dataset is reduced to 11 measurements. Furthermore, by checking the dust cases individually we observed that unlike the coarse mode, the fine mode is generally in good agreement between LIRIC and CAMx. In order to provide more information on the dust cases the old Figure 4 is modified and the dust cases are displayed with orange color.

Line 21: Please, specify the criteria you use to detect dust cases. Some trajectories do not seem to originate in dust source regions in Figure 1. Idem for continental.

To characterize the cases we check the trajectories separately in the PBL and the FT. Then, one trajectory, either in the PBL or in the FT, is required in order to identify the measurement as dust or biomass burning. Additionally, an empirical criterion of a maximum dust concentration above 10ugr/m^3 in the DREAM profile is also applied to ensure that the trajectory carries dust.

Line 33: This sentence is confusing. Rewrite. What is the diameter for separation between fine and mode in CAMx? Is it the same as in LIRIC?

See the relative comment response to Reviewer 1.

The text has been modified according to the reviewer's suggestions: "Another hindrance in the analysis is that the fine and coarse mode of LIRIC are not directly comparable with the PM2.5 and PM2.5-10 modes of CAMx. The PM2.5 particles should include all the fine particles and a small part of the coarse particles that changes depending on the case. Additionally, the size distribution of the sunphotometer usually surpasses the PM10 diameter limit. Fortunately, it is possible to convert the fine and coarse modes of LIRIC to PM2.5 and PM2.5-10 particles. In the LIRIC inversion, the normalized volume size distribution of each mode is derived by separating the columnar size distribution of the sunphotometer in the two modes. The normalized distribution of each mode remains constant with height. Taking that into account, the fractions of the sunphotometer's coarse mode that belong in the PM2.5 region and the region outside the PM10 particles can be calculated from the sunphotometer's volume size distribution. Then, the fine and coarse concentration profiles of each LIRIC case can be converted to PM2.5 and PM2.5-10 profiles using the equations 2 and 3. ”

Page 9

Line 1: Specify here the number of cases for the comparison. Why does this number emphasize the need of statistics?

The text was modified according to the reviewer’s suggestions: “A total of 22 cases take part in the comparison. We preferred a statistical approach in the analysis rather than comparing each case individually since the size of the dataset permits it.”

Lines 11-12: Provide more updated references.

The text was modified to: “(e.g., Flamant et al., 1997; Menut et al., 1999; Brooks, 2003; Tomasi and Perrone, 2006; Bravo-Aranda et al., 2016)

Line 15: Is this identification criteria based on a sensitivity analysis, previous studies, etc? Please, explain.

The reason why we use the upper limit criteria is that by applying the WTC method it is possible that a strong elevated layer could be identified as the PBL. This is also specified in Baars et al., 2008. In one of the cases they analyzed, an elevated dust layer complicated the derivation of the PBL top. Garrett, 1992 mention that the ABL typically extends from the ground to 2–3 km. Additionally, Georgoulias et al. 2009 in their study show that for noon measurements the mixing layer top is most of the time below 2600m for Thessaloniki. The selection of the upper limit value at 2500m is based on these studies.

The lower boundary corresponds to the height where the overlap function of the lidar system is above 0.9, that is 600m. The value of 400m is a typo and will be corrected. This is also mentioned in the response to Reviewer 1.

The text is modified according to the reviewer’s suggestion “Identification criteria are necessary for the selection of the PBL height. The top of the layer between 600m and 2.5km with the minimum value in the transformed signal is chosen as the boundary layer height. The upper limit is necessary in order to avoid identifying the top of sharp elevated layers as the PBL. According to Georgoulias et al. (2009) the upper limit of 2.5km is realistic for Thessaloniki. Baars et al. (2008) presented a case where an elevated dust layer complicated the PBL height retrieval with the WCT method. The wavelet transform is applied to the LIRIC concentration profiles before the upscaling of the resolution."

Line 17: What it is the advantage of applying the WCT to LIRIC output profiles instead of the range-corrected signal as in previous studies?

The range-corrected signal is an optical product that is typically used for the boundary layer height calculation because it is representative of the aerosol quantity and it is also much more straightforward to calculate than i.e. the aerosol backscatter or the aerosol extinction coefficient. Here, the aerosol concentration is already available and it provides direct information of the quantity of the aerosols. Thus, we preferred to use the LIRIC products instead.

Do you obtain similar results using the volume concentration profiles and the RCS?

The application of the WCT either in the range-corrected signal or in the LIRIC concentration provides similar results.

Line 23: No aerosol is expected above the upper limit in LIRIC, why don't you set these values to zero instead of a constant value?

The vertical profiles of CAMx extend up to 9.5km. As it can be seen from figures 3a and 3b the model typically provides non zero values in the whole profile so it will not be realistic to assume that the concentration is zero above the upper limit. Consequently, we use the information of the last point of the profile as the best guess of the aerosol load above that height. In figures 3a and 3b it can be seen that this choice produced a mean concentration of 1-2ugr/m^-3 at 10km for both the fine and the coarse particles which is not abnormal for this atmospheric region.
However, it is true that in rare occasions when the SNR in the lidar signals is quite low, especially near the upper limit, the LIRIC inversion can be affected. Higher than expected concentration values can be produced near the upper limit resulting in unrealistic LIRIC overestimations for the integrated mass values in the FT. We detected that this is the case for two measurements in the current dataset, one that belongs to the "continental" category and one that belongs in the "fires" category. In order to ensure the quality of our reference data we decided to remove those two cases from

the analysis. This reduces the total number of cases to 22, the number of the “fires” cases to 5 and the number of the “non fires” cases to 16.

Page 10

Line 17: More discussion, including numerical values, is missing here. For the case on January 13, 2014, it looks like most of the aerosol concentration is below the full overlap height. How does this affect the output profiles? How reliable are LIRIC output profiles in this case? Please, add some discussion in this respect.

See the relative comment response in Reviewer 1

Lines 19-26: As it is presented, it is not very clear what the contribution of the analysis of the optical properties to the evaluation is. Considering that the goal of the paper is the evaluation of CAMx, I think this section should be shorter or rewritten to clarify its purpose. Additionally, previous studies have shown that backscatter provided by LIRIC is affected by large uncertainties, especially for non-spherical particles (see Wagner et al., 2013 or Granados-Munoz et al., 2014). How do these backscatter profiles compared to those retrieved with a different method (e.g. Klett-Fernald)?

The agreement between LIRIC and Klett derived optical properties is very good. We present in the paper two typical cases as a demonstration of the methodology used to examine the aerosol profiles for each individual case. The inclusion of lidar ratio and angstrom exponent profiles provides further evidence for identifying different aerosol types and layers, but such profiles can be misleading if someone does not examine in parallel the extinction and backscatter profiles. A relevant comment has been added in the text

Add text

Line 20: specify if it’s extinction or backscatter related Angstrom exponent.

It’s the extinction Angstrom exponent. The text was modified according to the reviewer’s suggestions.

Line 20: It should be figure 2e instead of figure 2d. Include also the CAMx profile in Figure 2f

The figures have been renamed according to the reviewer’s suggestion.

The CAMx profile in figure 2f is not provided since it is biased in a similar way with figure 2d. This is specified in the text. The space inside this small figure is also limited. Furthermore, the main point in this section is to shortly demonstrate the capabilities and all the possible products of LIRIC for two different aerosol cases.

Page 11

Lines 10-15: Add numerical values in the discussion. In general in this section 4.1, add more discussion taking into account the uncertainties and shortcomings in LIRIC (and the model if provided by the modellers).

Section 4.1 was modified according to the reviewer's suggestions

Page 12

Line 24: Can you provide some information about the boundary layer height values obtained in the study?

The boundary layer height retrievals of the cases vary between 600m and 2500m without showing any strong pattern. A slight preference for PBL values in the range 1000-1500m can be observed. However, one has to take into account that the cases are not uniformly distributed either in the annual and daily cycle. Both of these variables highly affect the PBL height.

Besides, because of the incomplete overlap, LIRIC uncertainty in the PBL should be higher than in the troposphere. Take it into account when discussing the results.

As it was mentioned in the previous comment responses we will include in the text that the lidar signals are overlap corrected down to 600m since it was not clearly specified. Consequently, the signals can be trusted down to 600m. Indeed the missing part of the signal (0-600m) that is assumed to be constant can produce uncertainties in the retrieval. Munoz et al. 2014 have studied the uncertainty of the LIRIC retrieval using different parts of the signal that were not overlap corrected, and thus always underestimated, within acceptable overlap values (above 0.8). They found that the produced uncertainty is higher in the near range in terms of absolute values. This approach, however, includes both the uncertainty of the part of signals that is not overlap corrected and the uncertainty of the assumption of constant signals below the lower limit. For that reason, it is uncertain if the height variability that they observe applies to our case.

Page 14

Line 7: “are presented”

The text was modified according to the reviewer’s suggestion.

Page 15

Line 2: Provide more details on the results obtained removing the dust cases

The following text was added: “The comparison for the PM2.5 particles is actually affected in a negative way due to the limited number of measurements in the dataset.”

The dust cases in all the scatterplots will also be marked with an orange color.

Line 7: Do you have information about the relative humidity above Thessaloniki during the study period? This could give an idea about how important the hygroscopic growth is and how much it could affect the comparison. Consider rewriting the conclusions section after all previous comments.

Unfortunately, the only information available for this period is the water content which is added in the PM2.5 calculation. This could be analyzed in a future study.

Table 2: Should be a instead of z (or vice-versa)?

See the relative comment response to Reviewer 1

Figure 3: Add also the number of cases for the no fires category in the figure

Figure 3 has been updated according to the reviewer’s suggestions.

Investigating the quality of modeled aerosol profiles based on combined lidar and sunphotometer data

Nikolaos Siomos[1], Dimitris S. Balis[1], Anastasia Poupkou[1], Natalia Liora[1], Spyridon Dimopoulos[1], Dimitris Melas[1], Eleni Giannakaki[2,5], Maria Filioglou[1,2], Sara Basart[3], and Anatoli Chaikovsky[4]

[1]Laboratory of atmospheric physics, Physics Department, Aristotle University of Thessaloniki, Greece
[2]Finnish Meteorological Institute, Atmospheric Research Centre of Eastern Finland, Kuopio, Finland
[3]Earth Sciences Department, Barcelona Supercomputing Center, BSC-CNS, Barcelona, Spain
[4]Institute of Physics, National Academy of Science, Minsk, Belarus
[5]Department of Environmental Physics and Meteorology, Faculty of Physics, University of Athens, Greece

Correspondence to: N. Siomos (nsiomos@physics.auth.gr)

Abstract. In this study we present an evaluation of the Comprehensive Air Quality Model with extensions ~~CAMx~~ (CAMx) for Thessaloniki using radiometric and lidar data. The aerosol mass concentration profiles of CAMx are compared against the ~~fine and coarse mode aerosol~~ PM2.5 and PM2.5-10 concentration profiles retrieved by the Lidar-Radiometer Inversion Code LIRIC. The CAMx model and the LIRIC algorithm results were compared in terms of mean mass concentration profiles, center of mass and integrated mass concentration in the boundary layer and the free troposphere. The mean mass concentration comparison resulted in profiles within the same order of magnitude and similar vertical structure for the ~~fine~~ PM2.5 particles. The mean center of mass values are also close with a ~~fractional bias of 24.8%~~mean bias of 0.57 km. On the opposite side, ~~the coarse mode appears to be underestimated by the model below 4km and overestimated above~~there are larger differences for the PM2.5-10 mode both in the boundary layer and in the free troposphere. In order to grasp the reasons behind the discrepancies, we investigate the effect of aerosol ~~components~~ sources that are not properly included in the model's emission inventory and in the boundary conditions such as the wildfires and the desert dust component. The identification of the cases that are affected by wildfires is performed using wind backward trajectories from the Hybrid Single Particle Lagrangian Integrated Trajectory Model HYSPLIT in conjunction with satellite fire pixel data from the MODerate-resolution Imaging Spectroradiometer (MODIS) Terra and Aqua global monthly fire location product MCD14ML. By removing those cases the correlation coefficient improves from ~~0.44 to 0.86 for the fine mode~~ 0.69 to 0.87 for the PM2.5 integrated mass in the boundary layer ~~. The fine mode~~ and from 0.72 to 0.89 in the free troposphere. The PM2.5 center of mass fractional bias also decreases to ~~16.9%~~0.38 km. Concerning the analysis on the desert dust component, the simulations from the ~~updated version of the former~~ Dust Regional Atmospheric Model ~~called~~ (BSC-DREAM8b) were deployed. When only the ~~desert~~ Saharan dust cases are taken into account, BSC-DREAM8b generally outperforms CAMx when compared with LIRIC, achieving a correlation of 0.91 and a ~~fractional bias of -18.9~~mean bias of -29.1% for the integrated mass in the free troposphere and a correlation of ~~0.44~~ 0.57 for the center of mass. CAMx, on the other hand, ~~both underestimates and anti-correlates~~ underestimates the integrated mass in the free troposphere. Consequently, the accuracy of CAMx is limited concerning the transported Saharan dust cases. We conclude that the performance of CAMx appears to be best for the ~~fine particles, especially~~ PM2.5 particles, both in the boundary layer ~~. At~~

~~the same time it systematically fails to successfully predict the coarse mode~~and in the free troposphere. Sources of particles not properly taken into account by the model are confirmed to negatively affect its performance, especially for the PM2.5-10 particles.

1 Introduction

There is a wide variety of atmospheric models that are capable of providing vertical profiles of the aerosol mass concentration (e.g. CAMx, BSC-DREAM8b, ~~European Monitoring and Evaluation Programme EMEP,~~ LOng Term Ozone Simulation - EURopean Operational Smog model LOTOS-EUROS, CHIMERE). This is achieved through simulation of the atmospheric motion and the chemical reactions that are taking place inside the atmosphere. The most common approach to validate the modeled vertical mass concentration products is to compare with surface and columnar concentration or optical measurements either from ground-based or satellite instruments (e.g., Takemura et al., 2002; Stier et al., 2005; Katragkou et al., 2010; Huneeus et al., 2011; Basart et al., 2012a; Marécal et al., 2015). This approach, however, ~~doesn't~~ does not verify the ability of the model to accurately predict the vertical distribution of the aerosol concentration. Observational aerosol profiles comparable with the modeled ones are required for this purpose. Remote sensing techniques such as lidar measurements can provide us with this sort of profiles. Since the main lidar products typically involve optical aerosol properties such as the aerosol backscatter and extinction coefficient profiles, it is common to ensure comparability by converting the model's output after applying appropriate techniques. For example Mona et al. (2014) compare ~~between~~ the dust extinction profiles of the BSC-DREAM8b model and the respective EARLINET (European Aerosol Research LIdar NETwork) profiles for a 12 year period in Potenza. Meier et al. (2012) use lidar backscatter profiles as one of the tools to evaluate the Consortium for Small-scale Modeling - Multi-Scale Chemistry Aerosol Transport (COSMO-MUSCAT) model for the PM2.5 and PM10 particles. Hodzic et al. (2004) recreate the lidar attenuated backscatter profiles using the output of the CHIMERE model in order to compare the model's PM10 profiles with the lidar measurements.

On the other hand, there are techniques that allow the estimation of the aerosol vertical concentration from remote sensing lidar measurements using a suitable algorithmic inversion method (e.g., Böckmann, 2001; Veselovskii et al., 2002; Raut and Chazette, 2009; Lopatin et al., 2013; Chaikovsky et al., 2016). The advantage of this approach is that the modeled product can be directly validated without the need of conversion. The literature focused on the validation of dust transportation models with observational aerosol concentration profiles is quite mature. For example, Binietoglou et al. (2015) have presented a methodology based on LIRIC to evaluate the performance of dust models using data from multiple AERONET (AErosol RObotic NETwork) and EARLINET stations. Granados-Muñoz et al. (2016b) also use the LIRIC algorithm to compare between observational data and a variety of dust models in the frame of July 2012 CHemistry and AeRosols Mediterranean EXperiments ChArMEx/EMEP campaign. However, there is a lack of studies that focus on the evaluation of both PM2.5 and PM10 concentration profiles simulated by atmospheric models. Royer et al. (2011) compare the simulations of the CHIMERE chemistry transport model for the PM10 particles using lidar PM10 concentration profiles derived with the methodology of Raut and Chazette (2009). In this study we investigate the validity of the aerosol concentration profiles simulated with the air quality model Comprehensive Air Quality Model with extensions (CAMxversion5.3), (?) (ENVIRON 2010, 2010) for Thessaloniki, Greece (40.5N, 22.9E) using the results of the Lidar/Radiometer Inversion Code (LIRIC). Instead of evaluating the performance of CAMx only for the PM10 particles, we separate the fine from the coarse aerosols particles by applying the LIRIC technique, then we convert the fine and coarse concentration profiles of LIRIC to PM2.5 and PM2.5-10 profiles and perform the validation for each mode the PM2.5 and PM2.5-10 individually.

CAMx is running operationally to produce a 3-day air quality forecast for Thessaloniki (Zyryanov et al., 2012; Marécal et al., 2015). It provides vertical concentration profiles of a variety of gaseous and aerosol components.

A second model, the desert dust transportation model BSC-DREAM8b (Nickovic et al., 2001; Pérez et al., 2006a, b; Basart et al., 2012b) has also been included in the analysis in order to investigate the performance of CAMx in the case of dust transportation events. This model can provide total dust concentration profiles.

The LIRIC inversion (Chaikovsky et al., 2016), on the other hand, is a technique used to estimate the concentration profiles of the fine and coarse mode aerosol using both sunphotometer and lidar data. Lidar and sunphotometer measurements performed at the Laboratory of Atmospheric Physics of the Aristotle University of Thessaloniki, Greece (40.5N, 22.9E) from the period 2013-2015 were used as input data for the algorithm.

Validating the accuracy of CAMx simulations for Thessaloniki for the period 2013-2015 could prove useful in the aerosol classification procedure of the lidar measurements since individual aerosol components are provided by CAMx. Furthermore, from the modelers' point of view, the comparison could also reveal the need of adjustments in the model's aerosol emissions, boundary conditions and mixing processes.

The paper is organized as follows. In section 2 the two models, CAMx and BSC-DREAM8b, and the LIRIC algorithm are described in detail. The 3rd section is devoted to the methodology of the analysis. This includes the preprocessing pre-processing of the lidar and the sunphotometer measurements and the characterization of the lidar profiles, the demonstration of the strategy that we applied for the comparison and the application of two example cases. The results of the study are discussed in section 4. Finally, section 5 contains the main conclusions of this study.

2 Data, Algorithm and Models

2.1 The lidar system of Thessaloniki

Lidar measurements from the THEssaloniki LIdar SYSstem (THELISYS), that is located at the Laboratory of Atmospheric Physics (LAP) of the Aristotle University of Thessaloniki (40.5^o N, 22.9^o E) at 50m above sea level, during the period 2013-2015 were selected for this study. The setup of the system in this period includes two raman channels at 355nm and 532nm and three elastic channels at 355, 532 and 1064nm. The raw lidar signals from the elastic channels at the three aforementioned wavelengths are necessary in order to perform the LIRIC inversion. All signal pre-prepossessing procedures are applied directly in the LIRIC algorithm. The lidar station of Thessaloniki participates in the European Aerosol Research Lidar Network (EARLINET) Schneider et al. (2000) (Schneider et al., 2000; Pappalardo et al., 2014) since 2000. More details on the instrument can be found in Amiridis et al. (2005) and (Giannakaki et al., 2010).

2.2 The CIMEL sunphotometer

In order to apply the LIRIC inversion, sunphotometer observations are necessary. The CIMEL multiband sun-sky photometer of Thesaloniki was deployed for this purpose. The instrument participates in We used measurements from a CIMEL multiband sun-sky photometer which was installed in Thessaloniki in 2003 as part of the AERONET Global Network. It belongs to the Laboratory of Atmospheric Physics (LAP) and it is located in a distance less than 50m from the lidar instrument (see section 2.1) in the same altitude. It automatically performs direct solar irradiance and sky radiance measurements at 340, 380, 440, 500, 670, 870, and 1020nm. The data processing is performed automatically with the AERONET inversion algorithms (Dubovik and King, 2000; Dubovik et al., 2006). The level 2 Version 2 Inversion data during the period 2013-2015 were used in this study. The instrument and the AERONET infrastructure are described in detail in Holben et al. (1998).

2.3 The Comprehensive Air quality Model with extensions (CAMx)

An air quality forecast modeling system was set-up in the framework of the EU Monitoring Atmospheric Composition and Climate (MACC) project. It consists of the meteorological model Weather Research and Forecasting model (WRF version 3.5.1) described in Skamarock et al. (2008) and the photochemical model CAMx (version 5.3). It is designed to provide the air quality forecast in four nested grids covering Europe (30 km spatial resolution European grid covering also a part of Sahara desert), the Eastern Mediterranean (10 km spatial resolution grid) and the Greek urban centers Thessaloniki and Athens (2 km spatial resolution grids). A nesting technique is applied in order to increase the accuracy at the area of interest i.e. Thessaloniki. The domains of CAMx are presented in figure 1. A single grid point in Thessaloniki is chosen at (40.633 N, 22.956 E) for the model outputs processed in the present study to coincide with the lidar measurements. The model grids are configured in 17 vertical layers extending up to about 9.5 km above ground level (agl). The temporal resolution of CAMx outputs and consequently of the simulated profiles is one hour.

The aerosol particles are modeled using a static two-mode coarse/fine scheme in CAMx for the representation of the particle size distribution. Fine particles have an aerodynamic a diameter smaller than 2.5 µm while coarse particles have an aerodynamic a diameter higher than 2.5 µm and smaller than 10 µm. A total of 20 individual aerosol components plus the aerosol water content that is absorbed by the hygroscopic particles are provided by the model (see table 1). The aerosol aqueous inorganic chemistry is applied according to the RADM-AQ aqueous chemistry algorithm (Chang et al., 1987). The partitioning of the inorganic aerosol constituents between the gas and aerosol phases is performed using the ISORROPIA thermodynamic module (Nenes et al., 1998). The secondary organic aerosol (SOA) formation/partitioning was performed with the use of the SOAP scheme (Strader et al., 1999). SOA are formed by non-methane volatile organic compounds of anthropogenic and natural origin. Details on the CAMx aerosol components such as the hygroscopicity and the fine or coarse mode type can be found in table 1.

CAMx is applied with the use of gaseous and particulate anthropogenic and natural emissions. Particulate matter emissions from natural sources (windblown dust and sea salt aerosol) and biogenic volatile organic compounds from vegetation are estimated using the Natural Emission MOdel NEMO version 1 (Markakis et al., 2009; Poupkou et al., 2010; Markakis, 2010) driven by the WRF meteorology. The Model for the Spatial and Temporal Distribution of Emissions (MOSESS) (Markakis et al., 2013) was applied for the calculation of spatially and temporally disaggregated and chemically speciated anthropogenic emission data of the following pollutants: CO, NOx, SO2, NH3, NMVOC, PM10 and PM2.5. The anthropogenic emissions were estimated using either activity data with methodologies and emission factors of the EMEP/CORINAIR - CORe INventory AIR emission inventory guidebook (EEA, 2006) or the emission database of The Netherlands Organization (TNO) for the reference year 2007 (Kuenen et al., 2011). Anthropogenic emissions for the following sources are accounted for: energy production, central heating, industry, transportation, waste treatment and disposal, agricultural activities (i.e. biomass burning, fertilization), extraction and distribution of fossil fuels. It is important to mention that particle emissions due to dust resuspension from agricultural activities and road traffic as well as the wildfire emissions are not currently included in CAMx simulations. Saharan dust emissions are taken into account only indirectly in the CAMx chemical boundary conditions provided by the global forecast modeling systems Integrated Forecasting System IFS-MOZART until October 2014 and C-IFS afterwards (Flemming et al., 2009; Morcrette et al., 2009; Stein et al., 2012) in the framework of the MACC project. The model has been evaluated during the MACC-II project (Marécal et al., 2015) ~~and also with observations from the National Monitoring Network of the Air Quality (Melas et al., 2017)~~.

2.4 The BSC-DREAM8b model

The transported desert dust particles and the forest fire particles are common categories of aerosol components in Thessaloniki (Amiridis et al., 2005; Basart et al., 2009; Giannakaki et al., 2010). As it has already been stated, the setup of CAMx in Thessaloniki includes the desert dust component only from the global boundary conditions and it ~~doesn't~~ does not include wildfire emissions at all. Biases in those two aerosol components are expected to affect both aerosol modes since the desert dust particles are coarse dominant (Shettle and Fenn, 1979; d'Almeida, 1987) and the biomass burning particles are fine dominant (Tesche et al., 2009; Groß et al., 2013).

The desert dust component can be further analyzed by comparing LIRIC with a dust specialized model. The dust transportation model BSC-DREAM8b (Pérez et al., 2006a, b; Nickovic et al., 2012; Basart et al., 2012b) was chosen for the comparison. BSC-DREAM8b is managed by the Barcelona Supercomputer Center (BSC) and provides operational forecasts since May 2009, and is also participating in the northern Africa–MiddleEast–Europe (NA-ME-E) node of the World Meteorological Organization (WMO) Sand and Dust Storm Advisory and Assessment System (SDS-WAS) project. The BSC-DREAM8b model is embedded into the Eta/NCEP atmospheric model and solves the mass balance equation for dust, taking into account the different processes of the dust cycle (i.e., dust emission, transport and deposition). The updated version of the model includes 8 particle size bins (0.1–10 µm radius range) and dust-radiative feedback.

The BSC-DREAM8b model has been evaluated for longer periods over northern Africa and Europe (e.g., Jiménez-Guerrero et al., 2008; Pay et al., 2010; Pay et al., 2012; Basart et al., 2012b, a; Gama et al., 2015) and against experimental campaigns in source regions during the SAharan Mineral dUst experiMent SAMUM-1 (Haustein et al., 2009) and the Bodélé Dust Experiments (BoDEx, Todd et al., 2008). Furthermore, daily evaluation of BSC-DREAM8b with near-real-time observations is conducted at BSC. Currently, the daily operational model evaluation includes satellites (MODIS and MSG - Meteosat Second

Generation) and AERONET sun photometers. Some comparisons between lidar and forecast models profiles were performed in terms of aerosol vertical distribution for specific Saharan dust events in the Mediterranean Basin (e.g., Balis et al., 2004; Pérez et al., 2006a; Amiridis et al., 2009; Mona et al., 2012; Gobbi et al., 2013; Amiridis et al., 2013; Mona et al., 2014). In addition, Binietoglou et al. (2015) includes BSC-DREAM8b as one of the models that participate in their analysis validating its performance against LIRIC retrievals in 10 EARLINET/AERONET stations.

The present analysis includes the daily runs of BSC-DREAM8b. The initial state of dust concentration in the model was defined by the 24 h forecast from the previous-day model run. The Final Analyses of the National Centers of Environment Prediction (NCEP/FNL; at $1^o \times 1^o$) at 0UTC were used every 24 hours as initial conditions and boundary conditions at intervals of 6h. The model configuration used for the present study includes 24 Eta vertical layers extending up to approximately 15 km in the vertical. The resolution is set to 0.3^o in the horizontal. The temporal resolution of the simulations is 3 h. The domain of simulation covers northern Africa, the Middle East and Europe. It is worth mentioning that re-suspended wind-blown dust and the considered desert dust sources are limited to northern Africa and the Middle East (< 35^o N) in the BSC-DREAM8b model.

2.5 The LIdar-Radiometer Inversion Code LIRIC

The LIRIC algorithm utilizes both radiometric data that have been processed by the AERONET inversion algorithm (Dubovik and King, 2000; Dubovik et al., 2006) and also raw lidar signals at three wavelengths (355 nm, 532 nm and 1064 nm) in order to estimate the aerosol concentration profiles for the fine and coarse particles. The radiometric data used as input includes the aerosol size distribution, the aerosol volume concentration in the two modes (fine and coarse), the aerosol optical depth (AOD) and the single scattering albedo (SSA) of each mode, the complex refractive index in each wavelength, the sphericity and the aerosol phase function. The minimum of the aerosol size distribution in the 0.194 - 0.576 µm radius range is used as the size boundary that separates the fine from the coarse mode. In case that the particle depolarization ratio is also provided, it is possible for the algorithm to further separate the spherical coarse particles from the non spherical ones. Nevertheless, only the fine and coarse mode retrievals are taking part in this study due to the lack of depolarization ratio profiles.

The algorithm main products are the vertical volume concentration profiles in the two modes. In brief, the algorithm searches for the profile per mode that gives the best agreement between the actual data and the reconstructed data from the algorithm, also demanding a certain degree of vertical smoothness in the final product. The reconstructed data include the aerosol backscatter profiles in the three lidar wavelengths and the columnar volume concentration values per mode. A detailed description of LIRIC can be found in Chaikovsky et al. (2016).

The effects of multiple user defined uncertainties, such as the upper and lower limit heights of the profile and the algorithm's regularization parameters, on the final result has been studied by Granados-Muñoz et al. (2014) ~~while ? present a sensitivity study of the effect of the algorithm's input parameters to the final profiles~~and Filioglou et al. (2017) for selective case studies in Granada and Thessaloniki respectively. They agree that the parameter that produces the biggest uncertainties is the lower limit height of the profile. Furthermore, the LIRIC retrievals have already been evaluated for volcanic and desert dust particles by Wagner et al. (2013) showing that the inversion can be accurate for two quite different types of aerosol. The aerosol extinction products of LIRIC has also been compared against the respective products from the Generalized Aerosol Retrieval from Radiometer and Lidar Combined data (GARRLiC) algorithm and against the retrievals from raman lidar measurements (Bovchaliuk et al., 2016). Finally, LIRIC has also been validated against in-situ aircraft measurements (e.g., Granados-Muñoz et al., 2016a; Kokkalis Granados-Muñoz et al. (2016a) compared the LIRIC retrievals with airborn in-situ measurements and found a promising agreement with the differences between the two staying within the expected uncertainties. Kokkalis et al. (2017) analyzed a biomass burning case. Their comparison between the LIRIC retrievals and the aircraft measurements resulted in a good performance of the algorithm for the fine particles. As a result it can be used as an independent reliable tool for the validation of CAMx.

3 Methodology

The analysis is divided in two parts. Section 3.1 corresponds to the ~~preprocessing~~ pre-processing of the algorithm's and the model's estimates in order to calculate comparable final products. The ~~characterization procedure~~ aerosol type identification of the lidar profiles is also described there. The methodology of the comparison is included in section 3.2. Two sample cases are presented in section 3.3 aiming to give an example of the products that the algorithm and the models can provide and also to demonstrate typical problems that occur in the analysis.

3.1 ~~Preprocessing~~Pre-processing

In the first part of this section, the ~~preprocessing~~ pre-processing procedure of the lidar measurements is described. In the second part, we present the methodology that was applied in order to characterize the lidar profiles.

3.1.1 Lidar ~~preprocessing~~pre-processing

The LIRIC algorithm requires both the raw lidar signal resulting from the atmospheric elastic backscattering in 355nm, 532nm and 1064nm and the Version 2 Inversions from AERONET. Lidar measurements performed in Thessaloniki during the period 2013-2015 were used for this purpose. Before September 2012, the setup of the lidar system in Thessaloniki was lacking a 1064nm channel that is necessary for the LIRIC inversion. A manual cloud screening process was also applied in order to remove all the cloud affected lidar measurements since LIRIC is not designed for cloud layers. In addition, only daytime measurements were used since the sunphotometer only operates during daytime. It is important to mention that both instruments are located close to one another and at the same altitude above sea level (see section 2.1 and section 2.2).

The sunphotometer data are processed by the AERONET algorithms in order to calculate the necessary aerosol properties which are required as input for the algorithm. In our analysis, the closest AERONET inversion to the central time of the lidar measurement was selected for the LIRIC retrievals. Cases with an absolute time difference that exceeded 3 hours between the sunphotometer measurement time and the central time of the lidar measurement were excluded. The lidar signal ~~preprocessing~~ pre-processing is performed directly in LIRIC (Chaikovsky et al., 2016) and includes the averaging, smoothing, background correction and range correction procedures as well as the normalization of the lidar signals and also the selection of the lower and upper height boundaries in the signal where the LIRIC inversion is going to be performed. All the signals are adjusted to a common vertical resolution with a constant step of 15m. A lower height boundary has to be determined due to the overlap function of the lidar system. Operationally, we apply the method of Wandinger and Ansmann (2002) for the calculation of the overlap function and the full overlap height. In the current dataset the full overlap height was calculated at 900m. The lower boundary is set correction however cannot be trusted down to the ground (Wandinger and Ansmann, 2002). In this study, we apply the correction down to 600m where the overlap function is still above 90% and use this height as the lower boundary of the LIRIC inversion. Below this height the lidar signals are considered constant during the LIRIC inversion. The concentration retrievals are also kept constant below 600m. As it was mention in section 2.3, this can produce uncertainties since the radiometric data corresponds to the whole atmospheric column and most of the aerosol are usually located in the boundary layer, close to the ground. According to Granados-Muñoz et al. (2014) the selection of the lower limit is the main source of error. They estimate the maximum uncertainty due to such intrinsic errors at 33% and the overall profile error to stay below 15% most of the time. They also mention that a use of an overlap correction could reduce this uncertainty.

The upper boundary depends on the maximum height where aerosol exist in a significant quantityand , that is, a region where the lidar signal from the aerosol backscattering can no longer be separated from the noise. This height can vary depending on the atmospheric conditions. The output data of the algorithm includes vertical volume concentration profiles of the fine and coarse mode particles in ppbv. By summing adding the concentration in the two modes one can calculate the total aerosol concentration.

The vertical resolution of the LIRIC products and the model products is different so it was necessary to upscale LIRIC to the resolution of each model. This can affect the vertical structure of the profiles for individual cases but in a statistical analysis those effects will be smoothed (Binietoglou et al., 2015). The temporal resolution of CAMx forecasts is one hour while the temporal resolution of BSC-DREAM8b forecasts is three hours. Each of the retrieval algorithm's profiles were matched to the models' profile closest to the central lidar measurement. Since LIRIC derived concentration values are in ppbv units while both models' profiles are in $\mu\mathrm{g}\cdot\mathrm{m}^{-3}$ units, it was necessary to apply a unit conversion that also requires the aerosol density. Despite the aerosol component densities of CAMx being known, we preferred to convert ppbv to $\mu\mathrm{g}\cdot\mathrm{m}^{-3}$ (equation 1) since $\mu\mathrm{g}\cdot\mathrm{m}^{-3}$ is more widely used as a concentration unit.

$$c_{\mu\mathrm{g}\cdot\mathrm{m}^{-3}} = 10^3 \cdot c_{\mathrm{ppbv}} \cdot \rho_{\mathrm{g}\cdot\mathrm{cm}^{-3}} \quad (1)$$

Where ρ is the mean aerosol density. Typical density values of 1.5 and 2.6 $\mathrm{g}\cdot\mathrm{cm}^{-3}$ for the fine and coarse mode particles were used respectively (Bukowiecki et al., 2011; Schumann et al., 2011; Kokkalis et al., 2013).

Another hindrance in the analysis is that the fine and coarse mode of LIRIC are not directly comparable with the PM2.5 and PM2.5-10 modes of CAMx. The PM2.5 particles should include all the fine particles and a small part of the coarse particles that changes depending on the case. Additionally, the size distribution of the sunphotometer usually surpasses the PM10 diameter limit. Fortunately, it is possible to convert the fine and coarse modes of LIRIC to PM2.5 and PM2.5-10 particles. In the LIRIC inversion, the normalized volume size distribution of each mode is derived by separating the columnar size distribution of the sunphotometer in the two modes. The normalized distribution of each mode remains constant with height. Taking that into account, the fractions of the sunphotometer's coarse mode that belong in the PM2.5 region and the region outside the PM10 particles can be calculated from the sunphotometer's volume size distribution. Then, the fine and coarse concentration profiles of each LIRIC case can be converted to PM2.5 and PM2.5-10 profiles using the equations 2 and 3.

$$c_{PM_{2.5}}(z) = c_{fine}(z) + c_{coarse}(z) \cdot \frac{\int_{r_{f-c}}^{r_{PM_{2.5}}} \frac{dV}{dr} \cdot dr}{\int_{r_{f-c}}^{r_c} \frac{dV}{dr} \cdot dr} \tag{2}$$

$$c_{PM_{2.5-10}}(z) = c_{coarse}(z) - c_{coarse}(z) \cdot \frac{\int_{r_{f-c}}^{r_{PM_{2.5}}} \frac{dV}{dr} \cdot dr + \int_{r_{PM_{10}}}^{r_c} \frac{dV}{dr} \cdot dr}{\int_{r_{f-c}}^{r_c} \frac{dV}{dr} \cdot dr}) \tag{3}$$

Where $c_{fine}, c_{coarse}, c_{PM_{2.5}}, c_{PM_{2.5-10}}$ are the concentration profiles of LIRIC before and after the conversion and $\frac{dV}{dr}$ is the aerosol volume size distribution of the sunphotometer as a function of the aerosol radius. The radii $r_c, r_{f-c}, r_{PM_{2.5}}, r_{PM_{10}}$ are in μm units and correspond to the upper limit of the sunphotometer size distribution, the separator radius between the fine and the coarse mode of the sunphotometer (see section 2.3) which is different for each dataset case and the PM2.5 and the PM10 separator radii respectively.

3.1.2 Characterization of the lidar profiles

It was mentioned in section 2.3 that the emission inventory of CAMx lacks the biomass burning aerosol emissions from wildfires. Additionally, the desert dust emissions are taken into account only indirectly in the CAMx chemical boundary conditions (section 2.3). In order to examine the effect of those cases on the comparison we group the cases in four categories. The first one is the total of the cases that will be refer to as 'all'. The second one contains the cases identified as biomass burning wildfire aerosol and will be referred to as "fires" from now on. The aerosol characterization is performed using a combination of model simulations and satellite data. It is described in the next paragraph. When the category 'fires' is screened from the category 'all', the category 'non fires' is formed. It contains the continental and desert dust cases. Finally, the desert dust cases are also isolated and are included in the category 'dust'.

The backward trajectories from HYSPLIT in conjunction with fire pixel data from the MODIS Terra and Aqua Global Monthly Fire Location Product (MCD14ML) are used to identify the fire cases. The dust cases characterization also utilizes the HYSPLIT trajectories in combination with the BSC-DREAM8b profiles. A pair of 6-day back-trajectories, one arriving in the boundary layer region and another in the Free Troposphere, were used. The technique that was utilized in order to estimate the boundary layer height per case is described in section 3.2. A high fire spot density on a region where the air masses are passing near ground is applied as a criterion for the wildfire cases identification.

The trajectories for the continental, desert dust and biomass burning cases are presented in figure ~~1~~ 2 with blue (a and b), orange (c and d) and black color (e and f) respectively. The left column contains the air masses that arrive in the boundary layer, typically around 1km while the right column the ones that arrive in the Free Troposphere, usually ranging between 3 and 4km. Each trajectory is accompanied by the corresponding accumulated 6-day fire pixels.

3.2 Comparison Strategy

The first part of the evaluation of CAMx is based on the comparison of the aerosol concentration with the LIRIC estimates for the 'all' category (section 4.1). The effect of the wildfire cases on the results is also examined in this section. In the second part, the accuracy of CAMx in events of transported Saharan dust is investigated (section 4.2).

The concentration profiles of all the ~~fine~~ PM2.5 aerosol components (table 1) are summed to create the ~~fine mode~~ PM2.5 concentration profile of CAMx. The same applies to the ~~coarse~~ PM2.5-10 aerosol components of table 1. The water content is included in the ~~fine mode~~ PM2.5 group since all the hygroscopic particles are ~~fine~~in the same group.

~~The number of cases that~~ A total of 22 cases take part in the comparison~~as well as the high variability between individual cases emphasize the need of statistics in the analysis~~. We preferred a statistical approach in the analysis rather than comparing each case individually since the size of the dataset permits it. The mean profiles of the two models and LIRIC are calculated across the vertical range. The center of mass is also calculated for each case and each mode. It provides additional information on the height where the majority of the particles are located.

The Planetary ~~boundary layer~~ Boundary Layer (PBL) marks the top of the layer where the atmosphere is well mixed and the local aerosol component is also expected to be significant. On the other hand, the Free Troposphere (FT) above, is related to much less mixing and a stronger transported aerosol component is expected. Consequently, a comparison between CAMx and LIRIC in the boundary layer and in the free troposphere would be useful in order to investigate the accuracy of the model in these quite different atmospheric regions. Thus, we calculate the ~~fine and coarse~~ PM2.5 and PM2.5-10 integrated mass in the boundary layer and in the ~~Free Troposphere~~ free troposphere for each case.

There are multiple techniques in order to estimate the boundary layer from lidar measurements ~~(e.g., Flamant et al., 1997; Menut et al., 19~~ Baars et al. (2008) apply a wavelet covariance transform on the lidar range-corrected signal in order to translate signal layers to maxima and minima of the Wavelet Covariance Transform (WCT). Here we apply the transform to the total concentration derived by LIRIC. However this is not enough to automatically identify the PBL since the boundaries of multiple layers will be retrieved. Identification criteria are necessary for the selection of the PBL height. The top of the layer between ~~400m~~ 600m and 2.5km with the minimum value in the transformed signal is chosen as the boundary layer height. The upper limit is necessary in order to avoid identifying the top of sharp elevated layers as the PBL. According to Georgoulias et al. (2009) the upper limit of 2.5km is realistic for Thessaloniki. Baars et al. (2008) presented a case where an elevated dust layer complicated the PBL height retrieval with the WCT method. The wavelet transform is applied to the LIRIC concentration profiles before the upscaling of the resolution.

To perform the integration below and above the PBL a lower and an upper boundary is required. The LIRIC inversion requires a height where the aerosol content is not significant. This height is usually different for each case and typically ranges between 3km and 9km. On the other hand, CAMx always provides values up to 9.5km while BSC-DREAM8b provides values up to 15km. We used the ground level as the common lower boundary and the upper limit of CAMx (9.5km) as the common upper boundary. Since LIRIC profiles usually end below this upper limit, the last value of each profile is considered constant up to 9.5km.

Metrics for the center of mass and the integrated mass in the boundary layer and in the Free Troposphere are also calculated. This includes the mean values and the standard deviations for the algorithm and the model, the mean bias error, the mean fractional error the root mean square error (RMSE), the correlation coefficient and the least squares fit slope and axis intersect values. The equations for some of the metrics can be found on table 2.

In order to demonstrate how the comparison strategy was applied we present two distinct cases which includes the aerosol typing procedure, the concentration profiles comparison and the optical products that can be derived by the LIRIC algorithm.

3.3 Example Cases

The products of LIRIC, CAMx, BSC-DREAM8b and HYSPLIT for 2 case studies are presented here.

The main product of LIRIC is the ~~fine and coarse mode~~ PM2.5 and PM2.5-10 concentration profiles. Additionally, the aerosol extinction and backscattering efficiencies per mode and per wavelength (355nm, 532nm, 1064nm) are also derived ~~.~~ from the aerosol optical thickness E_f and E_c and the columnar concentration C_f and C_c in the fine and coarse mode respectively. The calculation of the extinction efficiencies is presented in the equations 4 and 5. The backscattering efficiencies are produced from the extinction efficiencies using the single scattering albedo and the phase function at 180^o. This procedure is described in (Chaikovsky et al., 2016).

$$Q_{ext,f}(\lambda) = \frac{E_f(\lambda)}{C_f(\lambda)} \tag{4}$$

$$Q_{ext,c}(\lambda) = \frac{E_c(\lambda)}{C_c(\lambda)} \tag{5}$$

The total extinction and backscatter coefficient profiles, symboled as a and b respectively, for the three wavelengths are calculated using the equations ~~2 and 3, were~~ 6 and 7, where Q_{ext} is the extinction efficiency and Q_{bsc} is the backscattering efficiency calculated by LIRIC. The concentration for the fine and coarse mode is marked as c_f and c_c respectively.

$$a(\lambda,z) = Q_{ext,f}(\lambda)\cdot c_f(z) + Q_{ext,c}(\lambda)\cdot c_c(z) \tag{6}$$

$$b(\lambda,z) = Q_{bsc,f}(\lambda)\cdot c_f(z) + Q_{bsc,c}(\lambda)\cdot c_c(z) \tag{7}$$

Then, from the LIRIC estimated extinction and backscatter profiles, the lidar ratio and the angstrom exponent can also be calculated. Furthermore, as it was mentioned in section 3.1, the concentration values below 600m are kept constant in the LIRIC inversion and this also applies to the optical products.

A typical continental case on the 13th of January 2014 and a typical dust case on the 27th of August 2013 were selected in order to demonstrate the comparison results for two quite different typical aerosol types. The ~~continental case~~ first is demonstrated on the left of figure ~~2 and the Saharan dust case~~ 3 and the latter on the right column.

The trajectories indicate a continental case on 13th of January 2014. They are presented in figure ~~2a~~3a. Additionally there could be some mixing with marine particles from the Adriatic sea. The air mass arriving in the free troposphere should be clean according to HYSPLIT since the trajectory is elevated, always above 3km. The concentration profiles of LIRIC and CAMx are presented in figure ~~2c and figure 2d~~3c and figure 3d. According to LIRIC the ~~fine mode is~~ PM2.5 particles are dominant below 1km. Above that height there is a ~~coarse~~ PM2.5-10 dominant layer which could be the result of mixing with marine aerosol. The ~~fine mode~~ PM2.5 aerosol concentration profile of CAMx seems in good agreement with LIRIC in the near range, below 2km. On the other hand, CAMx ~~fails to predict the coarse mode. Above 6km, an overestimation of CAMx can be observed in both profiles~~and LIRIC are not in a very good agreement for the PM2.5-10 particles.

The four optical products are also presented. The upper part of figure ~~2g~~ 3g panel contains the aerosol backscatter and extinction coefficient profiles and the lower part the lidar ratio and the extinction Angstrom exponent profiles. The lidar ratio values are around 70 sr^{-1} at 355nm and 65 sr^{-1} at 532nm below 1km. The Angstrom values are near 1.6 for the 355-532nm exponent and at 2.1 for the 532-1064nm exponent. Giannakaki et al. (2010) report a mean lidar ratio at 355nm of 56 $\pm$ 23 sr^{-1} and a backscatter-related Angstrom exponent at 355-532nm of 1.4 $\pm$ 1.0 for the continental polluted aerosol class in Thessaloniki. Between 1km and 2km the lidar ratio at 355nm drops to 55 sr^{-1} and the lidar ratio 532nm at 45 sr^{-1}. The Angstrom exponent in both regions drops near 1.3. According to Giannakaki et al. (2010) those values are still within the range of the acceptable range of the continental polluted class.

As far as the second case is ~~considered~~concerned, the trajectories indicate an event of transported Saharan dust in the FT (figure ~~2b~~3b). The air masses that arrive in the PBL seem to contain marine aerosol from the Mediterranean, probably mixed with emissions from local urban sources. A strong ~~coarse~~ PM2.5-10 mode can be observed in figure ~~2e~~ 3e both for the layer below 1.5km and the layer above. Despite CAMx predicting a promising ~~fine~~ PM2.5 mode (figure ~~2d~~3e), it ~~fails to do so for the coarse particles . This~~ is not in good agreement for the PM2.5-10 particles (not shown) and this is the reason why the BSC-DREAM8b concentration profile is presented here instead. This model describes the dust layer between 2km and 5km much better. ~~Below 2km the model isn't compatible any more with the observation~~This issue is discussed further in section 4.2.

As far as the optical products are concerned, the lidar ratio at 355nm and 532nm values range between 40 and 50 sr^{-1} for the whole profiles while the Angstrom exponent ranges between 1.0 and 1.5 for both spectral regions (figure ~~2h~~3h). Giannakaki et al. (2010) calculate a lidar ratio at 355nm of 52 $\pm$ 18 sr^{-1} and a backscatter-related Angstrom exponent at 355-532nm of 1.5 $\pm$ 1.0 for the Saharan dust aerosol class at Thessaloniki which seems compatible with this case. In the next section the statistical analysis is presented and the results are discussed.

4 Discussion

An ensemble of ~~24~~ 22 measurements, that fulfill the criteria described in section 3.2, took part in this comparison. These cases constitute the category 'all'. In the first part of this section (section 4.1) the comparison between LIRIC and CAMx is presented.

In the second part (section 4.2) the accuracy of CAMx in case of Saharan dust events is investigated using the simulations of BSC-DREAM8b.

4.1 Comparison between LIRIC and CAMx

In this section the simulated profiles of CAMx are compared against the observational profiles of LIRIC. The vertical mean profiles derived from LIRIC and CAMx for the fine and coarse mode PM2.5 and PM2.5-10 particles are displayed in figure 3a and figure 3b4a and figure 4b. The solid lines correspond to the 'all' category, while the dashed lines correspond to the 'no fires' category, which consists of 17 16 measurements. It can be seen that the fine mode PM2.5 mean concentration profiles show a very good agreement between LIRIC and CAMx. The vertical structure also seems to bear some similarities. There is an overall is similar. Below 2 km the model aerosol concentrations are up to 6-7 $\mu g \cdot m^{-3}$ lower compared to LIRIC while there is a slight overestimation by the model which becomes more significant above 2kmabove 3 km by 1-2 $\mu g \cdot m^{-3}$. Removing the 7 6 wildfire cases modifies the LIRIC mean profile towards slightly lower values and the CAMx mean profile towards slightly higher values, leading to smaller discrepancies below 1kmdiscrepancies smaller than 4 $\mu g \cdot m^{-3}$ for the whole profile. Details on the behavior of the model in the boundary layer and the free troposphere for both the 'all' and the 'no fires' categories can be found in the next paragraphs. As far as the coarse PM2.5-10 mode is considered, it seems to be severely underestimated by the model especially below 3km. Above 4.5km below 3 km. The concentration of CAMx stays below 5 $\mu g \cdot m^{-3}$ while LIRIC estimates a mean concentration around 20 $\mu g \cdot m^{-3}$. Above 4 km this behavior is reversed and the model seems to overestimate the aerosol concentration result in higher aerosol concentration providing values in the range 5-10 $\mu g \cdot m^{-3}$ while the concentration of LIRIC gradually drops below 2 $\mu g \cdot m^{-3}$. The reasons of these discrepancies the discrepancies in the PM2.5-10 mode could be connected with the model's emissions inventoryand , the boundary conditions or the transport processes within the simulation domain over long distances. Some insight can be offered by individually inspecting the contribution of each aerosol component in the final product as shown in the next paragraph.

The mean concentration profiles of the aerosol components that consist the fine and coarse mode PM2.5 and PM2.5-10 particles of CAMx are presented in figure 3c and figure 3d4c and figure 4d. The aerosol components of table 1 are grouped into categories that follow the OPAC formalism (Hess et al., 1998). The Particulate Chloride and Sodium components form the "sea salt" category. The Aerosol Water Content, the Primary Elemental Carbon, the Fine Crustal and Coarse Crustal as well as the Fine Other Primary and Coarse Other Primary components are independent and form the categories "water", "soot", "soil finePM2.5", "soil coarsePM2.5-10", "other finePM2.5" and "other coarsePM2.5-10" respectively. The particulate Nitrate, the Sulfate and the particulate Ammonium are all hygroscopic and are grouped into the "water soluble" category. The rest of the species are all organic and are grouped in the "organic insoluble" category. The Fine Other Primary component contains the fine PM2.5 particles that are considered as inert by the model not chemically speciated as well as a part of the fine PM2.5 sea salt that cannot be categorized neither as Particulate Chloride nor as Sodium. The Coarse Other Primary component contains the coarse particles that take part in chemical reactions PM2.5-10 particles that are chemically speciated like nitrate, sulfate, ammoniac, black elemental carbon and primary organic aerosol, as well as the coarse PM2.5-10 sea salt and all the other coarse PM2.5-10 particles that are considered as inert by the modelnot chemically speciated.

The dominant component below 2km is the ~~water soluble one~~"water soluble" one with maximum concentration values between 6-7 $\mu g \cdot m^{-3}$, followed by the ~~fine sea salt component and the water content~~ water content with a maximum value of 4 $\mu g \cdot m^{-3}$. Above 2km the ~~other fine~~ "soil PM2.5" component becomes significantly stronger with ~~a maximum~~ its maximum of 5 $\mu g \cdot m^{-3}$ located at 3-4km. ~~This region matches with the one where CAMx overestimates the fine mode. Consequently, the "other fine" component could be connected to these biases.~~ As far as the ~~coarse mode is~~ PM2.5-10 are considered, the majority of particles belong to the ~~soil coarse component~~ "other PM2.5-10" component. The "soil PM2.5-10" component is systematically lower than the "other PM2.5-10" component by at least 2 $\mu g \cdot m^{-3}$. Both components ~~seem~~ could be responsible for the ~~overestimation~~ higher values above 4km ~~in the coarse~~ compared to the LIRIC PM2.5-10 mean profiles since the highest values in both ~~both~~ component profiles are located ~~above 2km~~between 3 and 7 km. In order to investigate if only one of them or both are also responsible for the large bias below 3km we isolate the dust component by selecting only the dust cases. This comparison is presented in section 4.2. Below, the center of mass and the integrated mass are calculated in order to further quantify the comparison results.

The center of mass (table 2) provides information on the height where most of the particles are located. It is presented in LIRIC against CAMx center of mass scatter plots (figure ~~4~~5) for the ~~fine and the coarse mode~~PM2.5 and the PM2.5-10 particles. The least squares fit line and the correlation coefficient for the 'no fires' category is displayed in the plots. A synopsis of the center of mass metrics can be seen in table 3. The 'no fires' category is not included for the ~~coarse~~ PM2.5-10 aerosol. LIRIC estimates a mean center of mass value of ~~1.06~~ 2.43 $\pm$ ~~0.23~~ 0.76 km for the "all" category which increases slightly for the "no fires" category in the ~~fine~~ PM2.5 mode. CAMx predicts a somewhat higher mean center of mass at ~~1.36~~ 3.00 $\pm$ ~~0.47 km that doesn't~~ 0.83 km that does not change much if the wildfire cases are excluded. Thus, the resulting mean bias is ~~0.30~~ only 0.57 km and the fractional bias is ~~24.8~~20.9% and they improve to ~~0.20km and 16.9~~0.38 km and 14.3% respectively. The root mean square error (RMSE) ~~stays constant near 0.55km~~ is 0.81 km for the "no fires" category. Consequently, the height where most of the ~~fine~~ PM2.5 particles are located seems very well-predicted by the model. As far as the ~~coarse~~ PM2.5-10 mode is considered, LIRIC gives a mean value of ~~1.36~~ 1.93 $\pm$ ~~0.60~~ 0.76 km while CAMx predicts the center of mass at ~~3.16~~ 4.89 $\pm$ ~~1.21~~ 1.53 km. As a result, the mean bias and the RMSE are much larger here, which is expected given the ~~large~~ discrepancies that were discussed in the previous paragraph. The correlation coefficient for the ~~fine mode is 0.20 and increases to 0.34~~ PM2.5 particles is 0.19 and it increases to 0.46 when the wildfire cases are excluded. While the center of mass is useful when examining the vertical distribution and the location of the maximum concentration, it ~~doesn't~~ does not provide any insight on the concentration itself. Additional information on the accumulated concentration within an atmospheric region can be provided by calculating the integrated mass.

The comparison of the integrated mass in the boundary layer and the free troposphere is displayed in LIRIC against CAMx integrated mass scatterplots (figure ~~5~~6) and the metrics are presented in table 4. ~~The coarse mode isn't included since it seems to be significantly biased by the model.~~

The behavior of the model in the boundary layer is examined first. LIRIC estimates a mean integrated mass at ~~18.8~~ 23.6 $\pm$ ~~16.0~~ 18.9 $mg \cdot m^{-2}$ against a CAMx derived value of ~~14.1~~ 15.7 $\pm$ ~~10.5~~ 12.3 $mg \cdot m^{-2}$ for the "all" category which change to ~~14.1~~ 20.1 $\pm$ ~~8.6~~ 16.3 $mg \cdot m^{-2}$ and ~~16.1~~ 17.6 $\pm$ ~~10.5~~ 12.9 $mg \cdot m^{-2}$ respectively for the "no fires" category. The mean bias changes from ~~-4.7~~ -7.9 $\mathrm{mg \cdot m^{-2}}$ to ~~2.0~~ -2.5 $\mathrm{mg \cdot m^{-2}}$ that translates to a fractional bias ~~shift from -28.6 to 13.2~~improvement from -38.2 to -14.2%. The RMSE improves from ~~14.7~~ 13.7 $\mathrm{mg \cdot m^{-2}}$ to ~~6.1~~ 8.1 $\mathrm{mg \cdot m^{-2}}$. The correlation coefficient is at ~~0.44~~ 0.69 and the least square fit slope at ~~0.29~~ 0.45 for the "all" category but they both improve to ~~0.81 and 0.99~~ 0.87 and 0.69 respectively when the wildfire cases are removed. The results that occur for the "no fires" category indicate that, in the boundary layer, the lack of wildfire emissions in CAMx shouldn't be neglected~~as it obviously affect the performance of the model~~.

The behavior of the model in the free troposphere is ~~quite different. The LIRIC mean value is 18.8 ± 16.0 against a CAMx mean value of 14.1 ± 10.5 . By excluding the wild fire cases the LIRIC mean stays almost the same, at 23.1 ± 18.4 but the CAMx mean increases to 32.5 ± 31.9 . This causes the mean bias and RMSE to actually increase from 8.1 and 33.3 to 16.6 and 35.6 respectively. The fractional bias is also negatively affected and increases from 28.5 to 53.2%. On the other hand, the correlation coefficient is at 0.25 and the least square fit slope at 0.45 for the "all" category and they improve to 0.36 and 0.69 respectively for the "no fires" category. All things considered, the removal of the wildfire cases doesn't seem to positively affect the agreement between the CAMx and LIRIC in the free troposphere. This is also depicted in the scatterplot (figure 5, on the right) where the "fires" cases (red) are all quite close to the unity line in contrast with the boundary layer (figure 5, on the left) where most of those cases seem to be outliers~~also quite promising. The correlation of the model in the FT is also very high. The LIRIC mean values are in very good accordance with the mean values of CAMx. By excluding the wild fire cases the correlation improves from 0.72 to 0.89. Discrepancies in the free troposphere could also be attributed to the fact that the LIRIC profile above the height where the aerosol load is insignificant is considered constant (see sections 3.1.1 and 3.2) while the CAMx profile is still active in that region. ~~Additionally,~~ The effect of excluding the wildfire cases in the PBL comparisons is larger than in the FT ones which indicates a possible preference of the biomass burning layers to arrive in the PBL over Thessaloniki ~~could also be associated with this behavior~~rather than in the free troposphere.

Taking into account the very good performance of CAMx in both atmospheric regions, it appears to ~~predict the fine aerosol more accurately~~ produce somewhat higher values for the aerosol concentration in the boundary layer ~~than~~ and somewhat lower values in the free troposphere ~~. This behavior seems consistent with the previous results considering that CAMx tends to estimate higher concentration values in higher altitudes (figure 3 and figure 4) in the fine mode, which could also be attributed to the "other fine" component. When the wildfire cases are not taken into account the performance of the model improves but only in the PBL~~exhibiting a similar absolute fractional bias between 14-15% and a similar high correlation above 0.85 in both atmospheric regions for the PM2.5 particles. Possible causes for the discrepancies between LIRIC and CAMx, especially for the PM2.5-10 particles, could be the aerosol emission inventory of CAMx~~and~~ , the chemical boundary conditions and the long range aerosol transportation within the CAMx domain. The LIRIC inversion is also subject to uncertainties (see section 2.5) but their effect is certainly not enough to explain the large discrepancies observed for the PM2.5-10 particles. We have shown here that cases that also include wildfire aerosols are a challenge for the model since those emissions are not included at all. It has been stated in section 2 that the soil dust resuspension emissions are also not included and that the Saharan dust emissions are included ~~indirectly. The large~~ only as boundary conditions while the Saharan region is not completely outside the European domain (figure 1). The discrepancies in the ~~coarse mode~~ PM2.5-10 mean profile could be connected to ~~a combination of~~ the lack of dust resuspension emissions and ~~of biased Saharan dust boundary conditions~~Saharan dust emissions within the domain and also with issues associated with the boundary conditions or the transportation of desert dust over long distances. In the next section the behavior of CAMx in transported dust events is analyzed in more detail.

4.2 Dust cases analysis using BSC-DREAM8b

In this section the ~~coarse mode~~ products derived by LIRIC are compared against the simulations of both CAMx and BSC-DREAM8b. Out of the initial dataset of ~~24~~ 22 measurements, 6 were identified as dust cases. ~~We have to mention that the focus of this study is not a validation of BSC-DREAM8b since the number of dust cases wouldn't be sufficient.~~ The BSC-DREAM8b model has been previously extensively validated (see section 2.4) and it is used here solely to support the ~~analysis of CAMx~~discussion of the CAMx vs LIRIC comparisons for the PM2.5-10 particles. That being said, we aim to isolate the ~~coarse~~ desert dust component and compare between LIRIC and each model in order to investigate if the observed large discrepancies in the ~~coarse mode~~ PM2.5-10 products of CAMx are also present in the simulations of a model that specializes in desert dust. While it is feasible to get the ~~coarse~~ PM2.5-10 dust profile with CAMx from the ~~coarse~~ PM2.5-10 profile by selecting only the "soil ~~coarse~~PM2.5-10" component (table 1) this is not the case for ~~the coarse profile of~~ LIRIC. Having applied the aerosol characterization of section 3.1.2, it is reasonable to assume that the coarse mode, that also includes the PM2.5-10 particles, of the selected dust cases is almost entirely attributed to dust. Binietoglou et al. (2015) also use a dataset of measurements that were flagged as desert dust in order to compare between the observations and the simulations of the dust transportation models. ~~Obtaining the coarse~~ Isolating either the coarse or the PM2.5-10 dust of BSC-DREAM8b is also ~~not an option~~ challenging because this model provides total dust profiles. d'Almeida (1987) mention that the contribution of the fine dust should be low, especially near where it is emitted. Mamouri and Ansmann (2014) have separated the fine and coarse mode dust profiles during an outbreak event. They report that while the fine dust contribution in the mass concentration was significant near the ground in that transported dust case, it stayed well below 20% of the total concentration above 400m. Consequently, the use of the total dust profile of BSC-DREAM8b shouldn't jeopardize much the validity of the comparison with LIRIC, especially in the free troposphere.

In section 2.1.1 we describe a procedure to convert the coarse profiles of LIRIC to PM2.5-10 (equations 2 and 3). While this conversion is convenient for the comparison between LIRIC and CAMx this is not the case for the dust specialized model. As it was mentioned above, the retrievals of BSC-DREAM8b are comparable with the coarse profile of LIRIC for the desert dust cases. For this reason the coarse profile of LIRIC will be used in the comparison with BSC-DREAM8b while the PM2.5-10 profile of LIRIC will be used in the comparison with CAMx.

The mean profiles of the dust aerosol mass concentration ~~in the coarse mode~~ is presented in figure ~~6.~~ 7. The comparison between CAMx and LIRIC for the PM2.5-10 particles can be seen on the left while the comparison between BSC-DREAM8b and LIRIC for the coarse particles is presented on the right. It is obvious that CAMx underestimates the concentration by providing values that never exceed 10 $\mu g \cdot m^{-3}$ while the LIRIC mean values raise up to ~~45~~ 30 $\mu g \cdot m^{-3}$ and potentially much higher for selective cases. However, above 4km there seems to be some structural agreement between the two profiles and above 7km the small values of the LIRIC profile are successfully reproduced by CAMx. This results in a center of mass predicted by CAMx for the "soil PM2.5-10" around 4 km. Consequently, the overestimation above 4km in figure 4b, that was discussed in section 4.1, is probably associated mainly with the "other PM2.5-10" component since this behavior does not occur here. Despite the large negative bias this structural agreement results in a correlation coefficient of 0.63 in the FT between CAMx and LIRIC integrated mass.

On the other hand, BSC-DREAM8b mean values are close to the ones derived by LIRIC between 2km and 4km, ranging between 20 and 40 $\mu g \cdot m^{-3}$. Below 2km, even BSC-DREAM8b seems to underestimate the concentration. This could be attributed to mixing with ~~coarse~~ PM2.5-10 particles other than desert dust in this region. This scenario is further supported by taking into account the "dust" category trajectories that arrive in the PBL (figure ~~1c~~2c) in section 3.3. In the next paragraphs the center of mass and the integrated mass comparison is presented in a way similar to section 4.1.

The center of mass comparison between LIRIC and BSC-DREAM8b is presented in figure ~~7a~~ 8a accompanied by the observational and modeled metrics in table 5. ~~CAMx predicts a~~ The center of mass ~~at 2.55~~ value of BSC-DREAM8b is at 2.60 $\pm$ ~~1.13~~ 0.54 km which is ~~actually~~ quite close to the ~~value of BSC-DREAM8b at 2.38 $\pm$ 0.57. Discrepancies in the mean bias here originate mainly from differences in the LIRIC center of mass product that occur due to differences from the interpolation of the LIRIC profile to the vertical resolution of each model~~LIRIC estimates resulting in a mean bias of 0.03 km and a fractional bias of 0.2%. The correlation coefficient and the least square fit slope values are at ~~-0.13 and -0.28~~ 0.17 and 0.31 respectively between the algorithm and the air quality model but they improve at ~~0.44 and 0.39~~ 0.57 and 0.41 when the dust transportation model is used instead. Binietoglou et al. (2015) estimate a correlation coefficient of 0.38 for the same model and for a dataset of 69 dust profiles.

As far as the integrated mass is considered, the comparison of LIRIC vs BSC-DREAM8b is also presented in figure ~~7~~8, the boundary layer in figure ~~7b~~ 8b and the free tropospheric region in figure ~~7c~~8c. The agreement is best in the free tropospheric region for the dust specialized model which is in accordance with the profiles of figure ~~6.~~ 7. The mean values for BSC-DREAM8b in the FT are ~~90.1~~ 85.3 $\pm$ ~~94.9~~ 97.5 $mg \cdot m^{-2}$. They are quite close to the mean values of LIRIC at ~~108.9~~ 101.2 $\pm$ ~~72.9~~ 79.1 $mg \cdot m^{-2}$ resulting to the lowest absolute mean bias of table 6 at ~~-18.8~~ -15.9 $mg \cdot m^{-2}$ and a fractional error of ~~-18.9%. The respective fractional error of CAMx in the FT is -103.3~~-29.1%. The correlation between LIRIC and BSC-DREAM8b in the same region is high, at 0.91, and the least square fit slope is ~~1.19~~ 1.12 in contrast with the PBL region where the slope is much lower at 0.23 despite the correlation being similar. Binietoglou et al. (2015) report a correlation of 0.54 for the integrated dust concentration in the whole profile for their dataset of 69 measurements. As it has already been mentioned, discrepancies between the algorithm and the models can occur from the non dust aerosol component. This can be crucial for the PBL region, where particles from various sources are mixed.

The ~~comparison~~ comparisons between LIRIC and the two models ~~resulted to a much better performance of~~ highlight the differences between BSC-DREAM8b ~~than~~ and CAMx in the free troposphere for the dust cases. The dust specialized model was able to reproduce values similar to the observations, leading to the conclusion that the "soil ~~coarse~~PM2.5-10" component of CAMx is definitely underestimated, especially below 4km. This behavior could be linked to the model's lack of direct Saharan dust emissions within the domain (see section 2.3) as well as issues associated with the boundary conditions or transportation of dust over long distances. With the current dataset of 6 cases, however, it is not clear which of these mechanisms is the prevalent source of bias in the dust profiles of CAMx. However, we have to mention here that the two models are quite different concerning the parametrization of the size distribution of the soil dust and thus a direct quantitative comparison is not meaningful. Having said that, the desert dust cases are not the majority and by removing the 6 dust cases out of the ~~24 total cases didn't~~ 22 total cases did not improve much the ~~coarse~~ PM2.5-10 profile of CAMx. The comparison for the PM2.5 particles is actually affected in a negative way due to the limited number of measurements in the dataset. In order to explain these ~~large~~ biases (section 4.1) we have to assume that the "other ~~coarse~~PM2.5-10" component is also biased, most probably underestimated below 4km and overestimated above. This component contains the ~~coarse~~ PM2.5-10 sea salt and also the ~~coarse~~ PM2.5-10 ammoniac, sulfate and nitrate particles (see table 1 and section 4.1) that are all supposed to be hygroscopic. Despite that, the "other ~~coarse~~PM2.5-10" component is not treated as hygroscopic by the model. ~~Thus~~ Therefore, the water absorption and the hygroscopic growth of the ~~coarse particles is~~ PM2.5-10 particles are not taken into account at all. Consequently, the absence of any water content in the ~~coarse mode~~ PM2.5-10 particles of CAMx could lead to an underestimation of the model since LIRIC concentration profiles include the water content absorbed by the ~~fine and coarse~~ PM2.5 and PM2.5-10 particles.

5 Conclusions

The use of LIRIC in the evaluation of CAMx resulted in ~~a mean profile~~ very good quantitative and qualitative agreement for the PM2.5 aerosol profiles. The mean model profiles are in the same order of magnitude and ~~of~~ show a similar vertical distribution with the observational ~~one for the fine mode~~ones for the PM2.5 particles. The mean center of mass of the model is different by only ~~0.3~~0.57 km from the respective value of the ~~algorithm~~ LIRIC which translates to a mean fractional bias of ~~24.8~~20.9%. The ~~correlation coefficient is estimated at 0.20 and improves at 0.34 when the wildfires are excluded. The~~ integrated mass comparison indicates a ~~better~~ very good performance of the model both in the boundary layer ~~than~~ and in the free troposphere ~~. For the "no fires" category, the mean fractional bias improves from -28.6 to 13.2% . At the same time, the correlation coefficient of the integrated mass rises from 0.44 to 0.86 and the least squares fit slope from 0.29 to 0.99. The comparison in the free tropospheric region, on the other hand, is not clearly benefited from the removal of those cases~~with absolute fractional biases below 15% and correlation coefficients above 0.85 within the PBL and the FT. All things considered, there are strong indications that the lack of the wildfire emissions in CAMx ~~affect it's performance~~ affects the comparisons concerning the mass concentration of the ~~fine mode~~ PM2.5 particles that arrive in the boundary layer. When those cases are excluded the correlation between the model simulations and the observational data is very good, indicating that the model succeeds in representing most other PM2.5 sources and related processes.

The ~~coarse mode~~ PM2.5-10 mean profile of CAMx, on the other hand seems to be ~~greatly~~ underestimated below 4km and overestimated above. Consequently, the vertical structure ~~is also incompatible~~ shows differences with the observational data. ~~The~~ Concerning more specifically the behavior of the "soil ~~coarse~~PM2.5-10" component of CAMx~~was tested~~ , this was evaluated using selected dust cases and the desert dust dispersion model BSC-DREAM8b. ~~Both models underestimate the concentration~~

It was shown that CAMx underestimates the concentration by providing values that never exceed 10 $\mu\mathrm{g}\cdot\mathrm{m}^{-3}$ while the LIRIC mean values raise up to 30 $\mu\mathrm{g}\cdot\mathrm{m}^{-3}$ and potentially much higher for selective cases. This behavior is supposed to be linked to the model's lack of direct Saharan dust emissions within the domain as well as issues associated with the boundary conditions and the transportation of dust over long distances. The relative structural agreement observed between the two profiles result in a correlation coefficient of 0.63 in the boundary layer. In the free troposphere, BSC-DREAM8b achieves a fractional bias of -18.9% and a correlation coefficient at 0.91 in contrast with CAMx where the same metrics are estimated at -103.3% and -0.69 respectively. The center of mass is also better correlated in the dust model. Care has to be taken, however, because the interpolation of the LIRIC profiles in the models' vertical resolution affects the center of mass and shifts the LIRIC mean value by from 1.75km for the resolution of CAMx to 2.14km for the resolution of BSC-DREAM8b. Since FT between CAMx and LIRIC integrated mass. On the other hand, BSC-DREAM8b outperforms CAMx, at least in the free troposphere, it is reasonable to assume that the "soil coarse" component is a source of bias. Furthermore, the fact that mean values are close to the ones derived by LIRIC between 2km and 4km, ranging between 20 and 40 $\mu\mathrm{g}\cdot\mathrm{m}^{-3}$ and also show similar vertical structures.

Furthermore, since the small number of the dust cases (6 out of 2422) is not enough to explain the large discrepancies in the coarse mode PM2.5-10 mode between LIRIC and CAMx, it is likely that the "other coarsePM2.5-10" component is also biased. This could be linked to the fact that it consists of many subcomponents, some of which in theory are hygroscopic, like the coarse PM2.5-10 sea salt, but they are not treated as such by the model, possibly leading to underestimations in the aerosol concentration.

This study shows that the LIRIC code, based on the synergy of sunphotometer and lidar measurements can be used in order to evaluate an air quality model like CAMx as far as the aerosol mass concentration is considered. Furthermore, models specialized in particular types of emissions, like the BSC-DREAM8b dust transportation model, can be used along with LIRIC in order to help isolate one specific aerosol component that the air quality models provides or completely lacks. That way, the components can be tested individually, making it possible to directly associate biases with a specific type of emissions. HereIn this study, for example, we concluded that the lack of a wildfire component, the desert dust component and the remaining coarse PM2.5-10 component are all potential sources of bias in the modeled aerosol concentration profiles. The emissions that are associated with these aerosol types can then be examined and proper corrections could be applied in order to improve the overall performance of the model. Finally, if such comparisons are successful then the simulations of the model can also be utilized in the aerosol classification procedure of the lidar measurements since the individual aerosol components of the model could provide insight on the origin of the main aerosol layers.

Acknowledgements. The authors would like to acknowledge the EU projects MACC-III (Monitoring Atmospheric Composition and Climate - III, Grant agreement no: 633080) and MACC-II project (Monitoring Atmospheric Composition and Climate - Interim Implementation, Grant agreement no: 283576). The simulated results presented in this research paper have been produced using the EGI and HellasGrid infrastructures. The ACTRIS-2 project from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654109 is gratefully acknowledged. The authors would also like to acknowledge the support provided by the Scientific Computing Center at the Aristotle University of Thessaloniki throughout the progress of the work on air quality forecasting. BSC-DREAM8b simulations were performed on the Mare Nostrum supercomputer hosted by Barcelona Supercomputing Center-Centro Nacional de Supercomputación (BSC-CNS). S. Basart wants to acknowledge the CICYT project (CGL2013-46736). Elina Giannakaki acknowledges the support of the Academy of Finland (project no. 270108).

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, doi:10.1029/2005JD006190, d21203, 2005.

Amiridis, V., Kafatos, M., Perez, C., Kazadzis, S., Gerasopoulos, E., Mamouri, R. E., Papayannis, A., Kokkalis, P., Giannakaki, E., Basart, S., Daglis, I., and Zerefos, C.: The potential of the synergistic use of passive and active remote sensing measurements for the validation of a regional dust model, Annales Geophysicae, 27, 3155–3164, doi:10.5194/angeo-27-3155-2009, 2009.

Amiridis, V., Wandinger, U., Marinou, E., Giannakaki, E., Tsekeri, A., Basart, S., Kazadzis, S., Gkikas, A., Taylor, M., Baldasano, J., and Ansmann, A.: Optimizing CALIPSO Saharan dust retrievals, Atmospheric Chemistry and Physics, 13, 12 089–12 106, doi:10.5194/acp-13-12089-2013, 2013.

Baars, H., Ansmann, A., Engelmann, R., and Althausen, D.: Continuous monitoring of the boundary-layer top with lidar, Atmospheric Chemistry and Physics, 8, 7281–7296, doi:10.5194/acp-8-7281-2008, 2008.

Balis, D. S., 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, 31, doi:10.1029/2004GL019881, l13104, 2004.

Basart, S., Pérez, C., Cuevas, E., Baldasano, J. M., and Gobbi, G. P.: Aerosol characterization in Northern Africa, Northeastern Atlantic, mediterranean basin and middle east from direct-sun AERONET observations, Atmospheric Chemistry and Physics, 9, 8265–8282, 2009.

Basart, S., Pay, M. T., Jorba, O., Pérez, C., Jiménez-Guerrero, P., Schulz, M., and Baldasano, J. M.: Aerosols in the CALIOPE air quality modelling system: evaluation and analysis of PM levels, optical depths and chemical composition over Europe, Atmospheric Chemistry and Physics, 12, 3363–3392, doi:10.5194/acp-12-3363-2012, 2012a.

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, 2012b.

Binietoglou, I., Basart, S., Alados-Arboledas, L., Amiridis, V., Argyrouli, A., Baars, H., Baldasano, J. M., Balis, D., Belegante, L., Bravo-Aranda, J. A., Burlizzi, P., Carrasco, V., Chaikovsky, A., Comerón, A., D'Amico, G., Filioglou, M., Granados-Muñoz, M. J., Guerrero-Rascado, J. L., Ilic, L., Kokkalis, P., 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., Tsekeri, A., Vukovic, A., Wandinger, U., and Wagner, J.: A methodology for investigating dust model performance using synergistic EARLINET/AERONET dust concentration retrievals, Atmospheric Measurement Techniques, 8, 3577–3600, doi:10.5194/amt-8-3577-2015, 2015.

Böckmann, C.: Hybrid regularization method for the ill-posed inversion of multiwavelength lidar data in the retrieval of aerosol size distributions, Appl. Opt., 40, 1329–1342, doi:10.1364/AO.40.001329, 2001.

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

Bravo-Aranda, J. A., de Arruda-Moreira, G., Navas-Guzmán, F., Granados-Muñoz, M. J., Guerrero-Rascado, J. L., Pozo-Vázquez, D., Arbizu-Barrena, C., Olmo, F. J., Mallet, M., and Alados-Arboledas, L.: PBL height estimation based on lidar depolarisation measurements (POLARIS), Atmospheric Chemistry and Physics Discussions, 2016, 1–24, doi:10.5194/acp-2016-718, 2016.

Brooks, I. M.: Finding Boundary Layer Top: Application of a Wavelet Covariance Transform to Lidar Backscatter Profiles, Journal of Atmospheric and Oceanic Technology, 20, 1092–1105, doi:10.1175/1520-0426(2003)020<1092:FBLTAO>2.0.CO;2, 2003.

Bukowiecki, N., Zieger, P., Weingartner, E., Jurányi, Z., Gysel, M., Neininger, B., Schneider, B., Hueglin, C., Ulrich, A., Wichser, A., Henne, S., Brunner, D., Kaegi, R., Schwikowski, M., Tobler, L., Wienhold, F. G., Engel, I., Buchmann, B., Peter, T., and Baltensperger, U.: Ground-based and airborne in-situ measurements of the Eyjafjallajökull volcanic aerosol plume in Switzerland in spring 2010, Atmospheric Chemistry and Physics, 11, 10 011–10 030, doi:10.5194/acp-11-10011-2011, 2011.

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

Chang, J. S., Brost, R. A., Isaksen, I. S. A., Madronich, S., Middleton, P., Stockwell, W. R., and Walcek, C. J.: A three-dimensional Eulerian acid deposition model: Physical concepts and formulation, Journal of Geophysical Research: Atmospheres, 92, 14, doi:10.1029/JD092iD12p14681, 1987.

d'Almeida, G. A.: On the variability of desert aerosol radiative characteristics, Journal of Geophysical Research: Atmospheres, 92, 3017–3026, doi:10.1029/JD092iD03p03017, 1987.

Dubovik, O. and King, M. D.: A flexible inversion algorithm for retrieval of aerosol optical properties from Sun and sky radiance measurements, Journal of Geophysical Research: Atmospheres, 105, 20 673–20 696, doi:10.1029/2000JD900282, 2000.

Dubovik, O., Sinyuk, A., Lapyonok, T., Holben, B. N., Mishchenko, M., Yang, P., Eck, T. F., Volten, H., Muñoz, O., Veihelmann, B., van der Zande, W. J., Leon, J.-F., Sorokin, M., and Slutsker, I.: Application of spheroid models to account for aerosol particle nonsphericity in remote sensing of desert dust, Journal of Geophysical Research: Atmospheres, 111, doi:10.1029/2005JD006619, d11208, 2006.

EEA: EMEP/CORINAIR Emission Inventory Guidebook - 2006., Tech. Rep. 11/2006., European Environmental Agency, 2006.

ENVIRON 2010: User's Guide CAMx Comprehensive Air Quality Model with Extensions., Version 5.30, ENVIRON International Corporation, 2010.

Filioglou, M., Siomos, N., Poupkou, A., S., D., Chaikovsky, A., and Balis, D. S.: A sensitivity study of the LIdar-Radiometer Inversion Code (LIRIC) using selected cases from Thessaloniki database, International Journal of Remote Sensing, the article is currently under review, 2017.

Flamant, C., Pelon, J., Flamant, P. H., and Durand, P.: Lidar determination of the entrainment zone thickness at the top of the unstable marine atmospheric boundary layer, Boundary-Layer Meteorology, 83, 247–284, doi:10.1023/A:1000258318944, 1997.

Flemming, J., Inness, A., Flentje, H., Huijnen, V., Moinat, P., Schultz, M. G., and Stein, O.: Coupling global chemistry transport models to ECMWF's integrated forecast system, Geoscientific Model Development, 2, 253–265, doi:10.5194/gmd-2-253-2009, 2009.

Gama, C., Tchepel, O., Baldasano, J. M., Basart, S., Ferreira, J., Pio, C., Cardoso, J., and Borrego, C.: Seasonal patterns of Saharan dust over Cape Verde - a combined approach using observations and modelling, Tellus B, 67, 2015.

Georgoulias, A. K., Papanastasiou, D. K., Melas, D., Amiridis, V., and Alexandri, G.: Statistical analysis of boundary layer heights in a suburban environment, Meteorology and Atmospheric Physics, 104, 103–111, doi:10.1007/s00703-009-0021-z, 2009.

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, doi:10.5194/amt-3-569-2010, 2010.

Gobbi, G. P., Angelini, F., Barnaba, F., Costabile, F., Baldasano, J. M., Basart, S., Sozzi, R., and Bolignano, A.: Changes in particulate matter physical properties during Saharan advections over Rome (Italy): a four-year study, 2001–2004, Atmospheric Chemistry and Physics, 13, 7395–7404, doi:10.5194/acp-13-7395-2013, 2013.

Granados-Muñoz, M., Bravo-Aranda, J., Baumgardner, D., Guerrero-Rascado, J., Pérez-Ramírez, D., Navas-Guzmán, F., Veselovskii, I., Lyamani, H., Valenzuela, A., Olmo, F., Titos, G., Andrey, J., Chaikovsky, A., Dubovik, O., Gil-Ojeda, M., and Alados-Arboledas, L.: A comparative study of aerosol microphysical properties retrieved from ground-based remote sensing and aircraft in situ measurements during a Saharan dust event, Atmospheric Measurement Techniques, 9, 1113–1133, doi:10.5194/amt-9-1113-2016, cited By 3, 2016a.

Granados-Muñoz, M. J., Guerrero-Rascado, J. L., Bravo-Aranda, J. A., Navas-Guzmán, F., Valenzuela, A., Lyamani, H., Chaikovsky, A., Wandinger, U., Ansmann, A., Dubovik, O., Grudo, J. O., and Alados-Arboledas, L.: Retrieving aerosol microphysical properties by Lidar-Radiometer Inversion Code (LIRIC) for different aerosol types, Journal of Geophysical Research: Atmospheres, 119, 4836–4858, doi:10.1002/2013JD021116, 2013JD021116, 2014.

Granados-Muñoz, M. J., Navas-Guzmán, F., Luis Guerrero-Rascado, J., Antonio Bravo-Aranda, J., Binietoglou, 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., and Alados-Arboledas, L.: Profiling of aerosol microphysical properties at several EARLINET/AERONET sites during the July 2012 ChArMEx/EMEP campaign, Atmospheric Chemistry and Physics, 16, 7043–7066, 2016b.

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

Haustein, K., Pérez, C., Baldasano, J. M., Müller, D., Tesche, M., Schladitz, A., Esselborn, M., Weinzierl, B., Kandler, K., and von Hoyningen-Huene, W.: Regional dust model performance during SAMUM 2006, Geophysical Research Letters, 36, doi:10.1029/2008GL036463, l03812, 2009.

Hess, M., Koepke, P., and Schult, I.: Optical Properties of Aerosols and Clouds: The Software Package OPAC, Bulletin of the American Meteorological Society, 79, 831–844, doi:10.1175/1520-0477(1998)079<0831:OPOAAC>2.0.CO;2, 1998.

Hodzic, A., Chepfer, H., Vautard, R., Chazette, P., Beekmann, M., Bessagnet, B., Chatenet, B., Cuesta, J., Drobinski, P., Goloub, P., Haeffelin, M., and Morille, Y.: Comparison of aerosol chemistry transport model simulations with lidar and Sun photometer observations at a site near Paris, Journal of Geophysical Research: Atmospheres, 109, doi:10.1029/2004JD004735, d23201, 2004.

Holben, B., Eck, T., Slutsker, I., Tanré, D., Buis, J., Setzer, A., Vermote, E., Reagan, J., Kaufman, Y., Nakajima, T., Lavenu, F., Jankowiak, I., and Smirnov, A.: AERONET—A Federated Instrument Network and Data Archive for Aerosol Characterization, Remote Sensing of Environment, 66, 1 – 16, doi:http://dx.doi.org/10.1016/S0034-4257(98)00031-5, 1998.

Huneeus, N., Schulz, M., Balkanski, Y., Griesfeller, J., Prospero, J., Kinne, S., Bauer, S., Boucher, O., Chin, M., Dentener, F., Diehl, T., Easter, R., Fillmore, D., Ghan, S., Ginoux, P., Grini, A., Horowitz, L., Koch, D., Krol, M. C., Landing, W., Liu, X., Mahowald, N., Miller, R., Morcrette, J.-J., Myhre, G., Penner, J., Perlwitz, J., Stier, P., Takemura, T., and Zender, C. S.: Global dust model intercomparison in AeroCom phase I, Atmospheric Chemistry and Physics, 11, 7781–7816, doi:10.5194/acp-11-7781-2011, 2011.

Jiménez-Guerrero, P., Pérez, C., Jorba, O., and Baldasano, J. M.: Contribution of Saharan dust in an integrated air quality system and its on-line assessment, Geophysical Research Letters, 35, doi:10.1029/2007GL031580, l03814, 2008.

Katragkou, E., Zanis, P., Tegoulias, I., Melas, D., Kioutsioukis, I., Krüger, B. C., Huszar, P., Halenka, T., and Rauscher, S.: Decadal regional air quality simulations over Europe in present climate: near surface ozone sensitivity to external meteorological forcing, Atmospheric Chemistry and Physics, 10, 11 805–11 821, doi:10.5194/acp-10-11805-2010, 2010.

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

Kokkalis, P., Amiridis, V., Allan, J. D., Papayannis, A., Solomos, S., Binietoglou, I., Bougiatioti, A., Tsekeri, A., Nenes, A., Rosenberg, P. D., Marenco, F., Marinou, E., Vasilescu, J., Nicolae, D., Coe, H., Bacak, A., and Chaikovsky, A.: Validation of {LIRIC} aerosol concentration retrievals using airborne measurements during a biomass burning episode over Athens, Atmospheric Research, 183, 255 – 267, doi:http://dx.doi.org/10.1016/j.atmosres.2016.09.007, 2017.

Kuenen, J., van der Gon, H., Visschedijk, A., Dröge, R., and van Gijlswijk, R.: MACC European emission inventory for the years 2003-2007., Tech. Rep. TNO-060-UT-2011-00588, The Netherlands Organisation report, 2011.

Lopatin, A., Dubovik, O., Chaikovsky, A., Goloub, P., Lapyonok, T., Tanré, D., and Litvinov, P.: Enhancement of aerosol characterization using synergy of lidar and sun-photometer coincident observations: the GARRLiC algorithm, Atmospheric Measurement Techniques, 6, 2065–2088, doi:10.5194/amt-6-2065-2013, 2013.

Mamouri, R. E. and Ansmann, A.: Fine and coarse dust separation with polarization lidar, Atmospheric Measurement Techniques, 7, 3717–3735, doi:10.5194/amt-7-3717-2014, 2014.

Markakis, K.: Contribution to the study of particle pollution in Greece., Ph.D. thesis, Laboratory of Atmospheric Physics, Aristotle University of Thessaloniki., 2010.

Markakis, K., Giannaros, T., Poupkou, A., Liora, N., Melas, D., Sofiev, M., and Soares, J.: Evaluating the impact of particle emissions from natural sources in the Balkan region, in: European Aerosol Conference 2009, 6-9 September 2009, Karlsruhe, Germany, 2009.

Markakis, K., Katragkou, E., Poupkou, A., and Melas, D.: MOSESS: A New Emission Model for the Compilation of Model-Ready Emission Inventories-Application in a Coal Mining Area in Northern Greece, Environmental Modeling and Assessment, 18, 509–521, 2013.

Marécal, V., Peuch, V., Andersson, C., Andersson, S., Arteta, J., Beekmann, M., Benedictow, A., Bergström, R., Bessagnet, B., Cansado, A., Chéroux, F., Colette, A., Coman, A., Curier, R. L., Van Der Gon, H. A. C. D., Drouin, A., Elbern, H., Emili, E., Engelen, R. J., Eskes, H. J., Foret, G., Friese, E., Gauss, M., Giannaros, C., Guth, J., Joly, M., Jaumouillé, E., Josse, B., Kadygrov, N., Kaiser, J. W., Krajsek, K., Kuenen, J., Kumar, U., Liora, N., Lopez, E., Malherbe, L., Martinez, I., Melas, D., Meleux, F., Menut, L., Moinat, P., Morales, T., Parmentier, J., Piacentini, A., Plu, M., Poupkou, A., Queguiner, S., Robertson, L., Rouïl, L., Schaap, M., Segers, A., Sofiev, M., Tarasson, L., Thomas, M., Timmermans, R., Valdebenito, A., Van Velthoven, P., Van Versendaal, R., Vira, J., and Ung, A.: A regional air quality forecasting system over Europe: The MACC-II daily ensemble production, Geoscientific Model Development, 8, 2777–2813, 2015.

Meier, J., Tegen, I., Mattis, I., Wolke, R., Arboledas, L. A., Apituley, A., Balis, D., Barnaba, F., Chaikovsky, A., Sicard, M., Pappalardo, G., Pietruczuk, A., Stoyanov, D., Ravetta, F., and Rizi, V.: A regional model of European aerosol transport: Evaluation with sun photometer, lidar and air quality data, Atmospheric Environment, 47, 519 – 532, doi:http://dx.doi.org/10.1016/j.atmosenv.2011.09.029, 2012.

Melas, D., Liora, N., Poupkou, A., Giannaros, C., Dimopoulos, S., Kontos, S., Symeonidis, P., Progiou, A., Sitara, A., Economides, D., and Ziomas, I.: An Up-to-Date Assessment of the Air Quality in Greece with the Use of Modeling Tools, pp. 1123–1128, Springer International Publishing, Cham, doi:10.1007/978-3-319-35095-0_161, 2017.

Menut, L., Flamant, C., Pelon, J., and Flamant, P. H.: Urban boundary-layer height determination from lidar measurements over the Paris area, Appl. Opt., 38, 945–954, doi:10.1364/AO.38.000945, 1999.

Mona, L., Liu, Z., Mueller, D., Omar, A., Papayannis, A., Pappalardo, G., Sugimoto, N., and Vaughan, M.: Lidar measurements for desert dust characterization: an overview, Advances in Meteorology, 2012, doi:10.1155/2012/356265, copyright © 2012 L. Mona et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited, 2012.

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

Morcrette, J., Boucher, O., Jones, L., Salmond, D., Bechtold, P., Beljaars, A., Benedetti, A., Bonet, A., Kaiser, J. W., Razinger, M., Schulz, M., Serrar, S., Simmons, A. J., Sofiev, M., Suttie, M., Tompkins, A. M., and Untch, A.: Aerosol analysis and forecast in the european centre for medium-range weather forecasts integrated forecast system: Forward modeling, Journal of Geophysical Research Atmospheres, 114, 2009.

Nenes, A., Pandis, S. N., and Pilinis, C.: ISORROPIA: A new thermodynamic equilibrium model for multiphase multicomponent inorganic aerosols, Aquatic Geochemistry, 4, 123–152, 1998.

Nickovic, S., Kallos, G., Papadopoulos, A., and Kakaliagou, O.: A model for prediction of desert dust cycle in the atmosphere, Journal of Geophysical Research Atmospheres, 106, 18 113–18 129, 2001.

Nickovic, S., Vukovic, A., Vujadinovic, M., Djurdjevic, V., and Pejanovic, G.: Technical Note: High-resolution mineralogical database of dust-productive soils for atmospheric dust modeling, Atmospheric Chemistry and Physics, 12, 845–855, doi:10.5194/acp-12-845-2012, 2012.

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

Pay, M., Piot, M., Jorba, O., Gassó, S., Gonçalves, M., Basart, S., Dabdub, D., Jiménez-Guerrero, P., and Baldasano, J.: A full year evaluation of the CALIOPE-EU air quality modeling system over Europe for 2004, Atmospheric Environment, 44, 3322 – 3342, doi:http://dx.doi.org/10.1016/j.atmosenv.2010.05.040, 2010.

Pay, M. T., Jiménez-Guerrero, P., Jorba, O., Basart, S., Querol, X., Pandolfi, M., and Baldasano, J. M.: Spatio-temporal variability of concentrations and speciation of particulate matter across Spain in the CALIOPE modeling system, Atmospheric Environment, 46, 376–396, doi:10.1016/j.atmosenv.2011.09.049, 2012.

Poupkou, A., Giannaros, T., Markakis, K., Kioutsioukis, I., Curci, G., Melas, D., and Zerefos, C.: A model for European Biogenic Volatile Organic Compound emissions: Software development and first validation, Environmental Modelling and Software, 25, 1845–1856, 2010.

Pérez, C., Nickovic, S., Baldasano, J. M., Sicard, M., Rocadenbosch, F., and Cachorro, V. E.: A long Saharan dust event over the western Mediterranean: Lidar, Sun photometer observations, and regional dust modeling, Journal of Geophysical Research: Atmospheres, 111, doi:10.1029/2005JD006579, d15214, 2006a.

Pérez, C., Nickovic, S., Pejanovic, G., Baldasano, J. M., and Özsoy, E.: Interactive dust-radiation modeling: A step to improve weather forecasts, Journal of Geophysical Research: Atmospheres, 111, doi:10.1029/2005JD006717, d16206, 2006b.

Raut, J.-C. and Chazette, P.: Assessment of vertically-resolved PM_{10} from mobile lidar observations, Atmospheric Chemistry and Physics, 9, 8617–8638, doi:10.5194/acp-9-8617-2009, 2009.

Royer, P., Chazette, P., Sartelet, K., Zhang, Q. J., Beekmann, M., and Raut, J.-C.: Comparison of lidar-derived PM_{10} with regional modeling and ground-based observations in the frame of MEGAPOLI experiment, Atmospheric Chemistry and Physics, 11, 10 705–10 726, doi:10.5194/acp-11-10705-2011, 2011.

Schneider, J., Balis, D., Böckmann, C., Bösenberg, J., Calpini, B., Chaikovsky, A., Comeron, A., Flamant, P., Freudenthaler, V., Hågård, A., Mattis, I., Mitev, V., Papayannis, A., Pappalardo, G., Pelon, J., Perrone, M., Resendes, D., Spinelli, N., Trickl, T., Vaughan, G., and Visconti, G.: European aerosol research lidar network to establish an aerosol climatology (EARLINET), Journal of Aerosol Science, 31, S592–S593, cited By 15, 2000.

Schumann, U., Weinzierl, B., Reitebuch, O., Schlager, H., Minikin, A., Forster, C., Baumann, R., Sailer, T., Graf, K., Mannstein, H., Voigt, C., Rahm, S., Simmet, R., Scheibe, M., Lichtenstern, M., Stock, P., Rüba, H., Schäuble, D., Tafferner, A., Rautenhaus, M., Gerz, T., Ziereis, H., Krautstrunk, M., Mallaun, C., Gayet, J.-F., Lieke, K., Kandler, K., Ebert, M., Weinbruch, S., Stohl, A., Gasteiger, J., Groß, S., Freudenthaler, V., Wiegner, M., Ansmann, A., Tesche, M., Olafsson, H., and Sturm, K.: Airborne observations of the Eyjafjalla volcano ash cloud over Europe during air space closure in April and May 2010, Atmospheric Chemistry and Physics, 11, 2245–2279, doi:10.5194/acp-11-2245-2011, 2011.

Shettle, E. P. and Fenn, R. W.: Models for the aerosols of the lower atmosphere and the effects of humidity variations on their optical properties, Tech. rep., 1979.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, M., Duda, K. G., Huang, X. Y., Wang, W., and Powers, J. G.: A description of the Advanced Research WRF Version 3, Tech. rep., National Center for Atmospheric Research, 2008.

Stein, O., Flemming, J., Inness, A., Kaiser, J. W., and Schultz, M. G.: Global reactive gases forecasts and reanalysis in the MACC project, Journal of Integrative Environmental Sciences, 9, 57–70, 2012.

Stier, P., Feichter, J., Kinne, S., Kloster, S., Vignati, E., Wilson, J., Ganzeveld, L., Tegen, I., Werner, M., Balkanski, Y., Schulz, M., Boucher, O., Minikin, A., and Petzold, A.: The aerosol-climate model ECHAM5-HAM, Atmospheric Chemistry and Physics, 5, 1125–1156, doi:10.5194/acp-5-1125-2005, 2005.

Strader, R., Lurmann, F., and Pandis, S. N.: Evaluation of secondary organic aerosol formation in winter, Atmospheric Environment, 33, 4849–4863, 1999.

Takemura, T., Nakajima, T., Dubovik, O., Holben, B. N., and Kinne, S.: Single-Scattering Albedo and Radiative Forcing of Various Aerosol Species with a Global Three-Dimensional Model, Journal of Climate, 15, 333–352, doi:10.1175/1520-0442(2002)015<0333:SSAARF>2.0.CO;2, 2002.

Tesche, M., Ansmann, A., Müller, D., Althausen, D., Engelmann, R., Freudenthaler, V., and Groß, S.: Vertically resolved separation of dust and smoke over Cape Verde using multiwavelength Raman and polarization lidars during Saharan Mineral Dust Experiment 2008, Journal of Geophysical Research: Atmospheres, 114, doi:10.1029/2009JD011862, d13202, 2009.

Todd, M. C., Bou Karam, D., Cavazos, C., Bouet, C., Heinold, B., Baldasano, J. M., Cautenet, G., Koren, I., Perez, C., Solmon, F., Tegen, I., Tulet, P., Washington, R., and Zakey, A.: Quantifying uncertainty in estimates of mineral dust flux: An intercomparison of model performance over the Bodélé Depression, northern Chad, Journal of Geophysical Research: Atmospheres, 113, doi:10.1029/2008JD010476, d24107, 2008.

Tomasi, F. D. and Perrone, M. R.: {PBL} and dust layer seasonal evolution by lidar and radiosounding measurements over a peninsular site, Atmospheric Research, 80, 86 – 103, doi:http://dx.doi.org/10.1016/j.atmosres.2005.06.010, 2006.

Veselovskii, I., Kolgotin, A., Griaznov, V., Müller, D., Wandinger, U., and Whiteman, D. N.: Inversion with regularization for the retrieval of tropospheric aerosol parameters from multiwavelength lidar sounding, Appl. Opt., 41, 3685–3699, doi:10.1364/AO.41.003685, 2002.

Wagner, J., Ansmann, A., Wandinger, U., Seifert, P., Schwarz, A., Tesche, M., Chaikovsky, A., and Dubovik, O.: Evaluation of the Lidar/Radiometer Inversion Code (LIRIC) to determine microphysical properties of volcanic and desert dust, Atmospheric Measurement Techniques, 6, 1707–1724, doi:10.5194/amt-6-1707-2013, 2013.

Wandinger, U. and Ansmann, A.: Experimental determination of the lidar overlap profile with Raman lidar, Appl. Opt., 41, 511–514, doi:10.1364/AO.41.000511, 2002.

Zyryanov, D., Foret, G., Eremenko, M., Beekmann, M., Cammas, J. ., D'Isidoro, M., Elbern, H., Flemming, J., Friese, E., Kioutsioutkis, I., Maurizi, A., Melas, D., Meleux, F., Menut, L., Moinat, P., Peuch, V. ., Poupkou, A., Razinger, M., Schultz, M., Stein, O., Suttie, A. M., Valdebenito, A., Zerefos, C., Dufour, G., Bergametti, G., and Flaud, J. .: 3-D evaluation of tropospheric ozone simulations by an ensemble of regional Chemistry Transport Model, Atmospheric Chemistry and Physics, 12, 3219–3240, 2012.

Table 1. CAMx Aerosol Components Synopsis. The Fine Other Primary is mainly consisted out of the ~~fine~~ PM2.5 aerosols that are treated as inert by the model. It also contains small part of the ~~fine~~ PM2.5 sea salt that cannot be treated as Particulate Chloride or as Sodium. The Coarse Other Primary component includes all the ~~coarse~~ PM2.5-10 aerosols that are not crustal (nitrate, sulphate, ammoniac, black carbon and primary organic aerosols) as well as the ~~coarse~~ PM2.5-10 sea salt and all the other particles that are considered as inert by the model.

Components	Hygroscopic	Mode	Category
Particulate Nitrate (NO3)	yes	~~fine~~ PM2.5	water soluble
Sulfate (SO4)	yes	~~fine~~ PM2.5	water soluble
Particulate Amonium (NH4)	yes	~~fine~~ PM2.5	water soluble
Aerosol Water Content (H2O)	–	~~fine~~ PM2.5	water
Anthropogenic SOA*	no	~~fine~~ PM2.5	organic insoluble
Biogenic SOA*	no	~~fine~~ PM2.5	organic insoluble
Polymerized Anthropogenic SOA*	no	~~fine~~ PM2.5	organic insoluble
Polymerized Biogenic SOA*	no	~~fine~~ PM2.5	organic insoluble
Sodium (Na)	yes	~~fine~~ PM2.5	sea salt ~~fine~~PM2.5
Particulate Chloride (Cl)	yes	~~fine~~ PM2.5	sea salt ~~fine~~PM2.5
Primary Organic Aerosol	no	~~fine~~ PM2.5	organic insoluble
Primary Elemental Carbon (C)	no	~~fine~~ PM2.5	soot
Fine Other Primary	no	~~fine~~ PM2.5	other ~~fine~~PM2.5
Fine Crustal	no	~~fine~~ PM2.5	soil ~~fine~~PM2.5
Coarse Other Primary	no	~~coarse~~ PM2.5-10	other ~~coarse~~PM2.5-10
Coarse Crustal	no	~~coarse~~ PM2.5-10	soil ~~coarse~~PM2.5-10

*SOA = Secondary Organic Aerosol

Table 2. Common Metrics Equations. The model data is defined as M and the lidar observational data as O. The ~~a~~ z and c symbols correspond to the height and the concentration respectively.

Metric	Equation
Center of Mass	$\frac{\int z \cdot c \cdot dz}{\int c \cdot dz}$
Mean Bias	$\frac{1}{N}\sum_{i=1}^{N}(M_i - O_i)$
Mean Fractional Bias (%)	$\frac{200}{N}\sum_{i=1}^{N}\frac{(M_i - O_i)}{(M_i + O_i)}$
RMSE	$[\frac{1}{N}\sum_{i=1}^{N}(M_i - O_i)^2]^{\frac{1}{2}}$

Table 3. Center of Mass Metrics. The category 'All' corresponds to the total of the cases (~~24~~ 22 in total) while the category 'No Fires' refers to the cases that are not classified as wildfires (~~17~~ 16 out of ~~24~~22). The mean, standard deviation, mean bias and root mean square error values are in [km] units. The Pearson's correlation coefficient (r) and the least square fit slope (a) and intercept (b) values are also calculated.

Mode	Category	LIRIC Mean	CAMx Mean	Mean Bias	Frac. Bias (%)	RMSE	r	a
~~Fine~~ PM2.5	All	~~1.06~~ 2.43 ± ~~0.23~~ 0.76	~~1.36~~ 3.00 ± ~~0.47~~ 0.83	~~0.30~~ 0.57	~~24.8~~ 20.9	~~0.56~~ 1.02	~~0.20~~ 0.19	~~0.23~~ 0
~~Fine~~ PM2.5	No Fires	~~1.14~~ 2.56 ± ~~0.44~~ 0.76	~~1.35~~ 2.94 ± ~~0.50~~ 0.80	~~0.21~~ 0.38	~~16.9~~ 14.3	~~0.54~~ 0.81	~~0.34~~ 0.46	~~0.39~~ 0
~~Coarse~~ PM2.5-10	All	~~1.36~~ 1.93 ± ~~0.60~~ 0.76	~~3.16~~ 4.89 ± ~~1.21~~ 1.10	~~1.80~~ 2.96	~~79.6~~ 85.8	~~1.31~~ 1.53	~~0.07~~ -0.34	~~0.13~~ -

Table 4. Fine Mode PM2.5 Integrated Mass Metrics. Two atmospheric regions are provided. The region below the boundary layer is symbolized as PBL while the region above it, in the free troposphere, is defined as FT. The category 'All' corresponds to the total of the cases (24 22 in total) while the category 'No Fires' refers to the cases that are not classified as wildfires (17 16 out of 2422). The mean, standard deviation, mean bias and root mean square error values are in $[\mathbf{mg} \cdot \mathbf{m}^{-2}]$ units. The Pearson's correlation coefficient (r) and the least square fit slope (a) and intercept (b) values are also calculated.

Region	Category	LIRIC Mean	CAMx Mean	Mean Bias	Frac. Bias (%)	RMSE	r	a	
PBL	All	18.8 23.6 ± 16.0 18.9	14.1 15.7 ± 10.5 12.3	-4.7 -7.9	-28.6 -38.2	14.7 13.7	0.44 0.69	0.29 0.45	8.61
PBL	No Fires	14.1 20.1 ± 8.6 16.3	16.1 17.6 ± 10.5 12.9	2.0 -2.5	13.2 -14.2	6.1 8.1	0.81 0.87	0.99 0.69	2.16
FT	All	24.4 30.9 ± 17.4 22.3	32.5 33.6 ± 31.9 33.4	8.1 2.7	28.5 2.0	33.3 23.2	0.25 0.72	0.45 1.08	21.4
FT	No Fires	23.2 31.0 ± 18.4 22.8	40.0 41.0 ± 35.4 36.7	16.8 10.1	53.2 14.9	35.6 19.4	0.36 0.89	0.69 1.43	24.0

Table 5. Dust Center of Mass Metrics. Metrics for the two models and LIRIC are provided. A total of 6 out of ~~24~~ 22 cases are included in the category 'dust'. Differences in the LIRIC values per model are attributed to the interpolation in different vertical resolution depending on the model. The mean, standard deviation, mean bias and root mean square error values are in $[\mathbf{km}]$ units. The Pearson's correlation coefficient (r) and the least square fit slope (a) and intercept (b) values are also calculated.

Model	LIRIC Mode	LIRIC Mean	Model Mean	Mean Bias	Frac. Bias (%)	RMSE
CAMx	~~1.75 ± 0.51~~ PM2.5-10	2.55 ± ~~1.13~~ 0.70	~~0.79~~ 4.04 ± 1.28	~~36.7~~ 1.48	~~1.30~~ 41.6	~~-0.13~~ 1.35
~~BSC-DREAM8b~~ DREAM	~~2.14~~ Coarse	2.63 ± ~~0.63~~ 0.75	~~2.38~~ 2.60 ± ~~0.57~~ 0.54	~~0.23~~ -0.03	~~10.2~~ 0.2	~~0.64~~ 0.63

Table 6. Dust Integrated Mass Metrics. Metrics for the two models in two atmospheric regions are provided. The region below the boundary layer is symbolized as PBL while the region above it, in the free troposphere, is defined as FT. A total of 6 out of ~~24~~ 22 cases are included in the category 'dust'. The mean, standard deviation, mean bias and root mean square error values are in $[\mathbf{mg \cdot m^{-2}}]$ units. The Pearson's correlation coefficient (r) and the least square fit slope (a) and intercept (b) values are also calculated.

Model	Region	LIRIC Mode	LIRIC Mean	Model Mean	Mean Bias	Frac. Bias (%)	RMSE
CAMx	PBL	~~39.7~~ PM2.5-10	34.7 ± ~~50.9~~ 38.0	~~3.3~~ 0.8 ± ~~2.2~~ 0.9	~~-36.4~~ -33.8	~~-169.3~~ -189.5	~~50.2~~ 37.6
~~BSC-DREAM8b~~ DREAM	PBL	~~43.8~~ Coarse	54.2 ± ~~52.6~~ 57.7	~~10.8~~ 14.5 ± ~~13.2~~ 15.3	~~-33.0~~ -39.7	~~-120.9~~ -114.5	~~40.6~~ 45.2
CAMx	FT	~~106.0~~ PM2.5-10	65.4 ± ~~74.9~~ 56.8	~~33.8~~ 18.4 ± ~~23.8~~ 26.9	~~-72.2~~ -47.0	~~-103.3~~ -139.0	~~92.9~~ 44.9
~~BSC-DREAM8b~~ DREAM	FT	~~108.9~~ Coarse	101.2 ± ~~72.9~~ 79.1	~~90.1~~ 85.3 ± ~~94.9~~ 97.5	~~-18.8~~ -15.9	~~-18.9~~ -29.1	~~41.3~~ 42.0

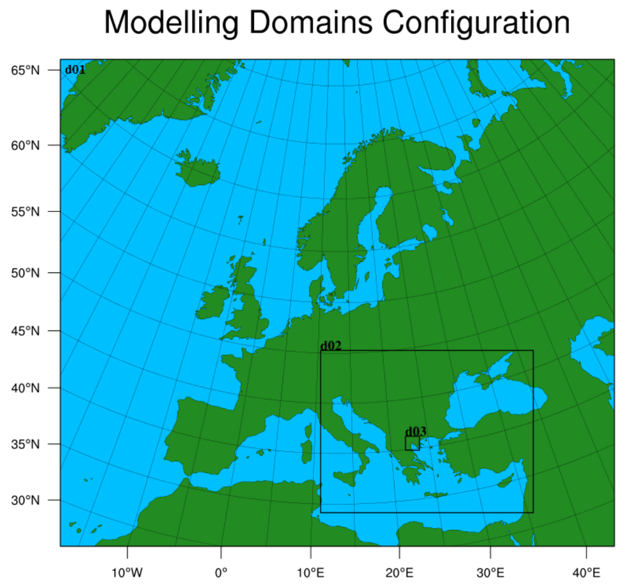


Figure 1. The three domains of CAMx, the European domain (d01), the Balkan domain (d02) and the domain of Thessaloniki (d03). A nesting technique is applied in order to increase the accuracy in the inner domains.

6 Days Backward Trajectories over Thessaloniki

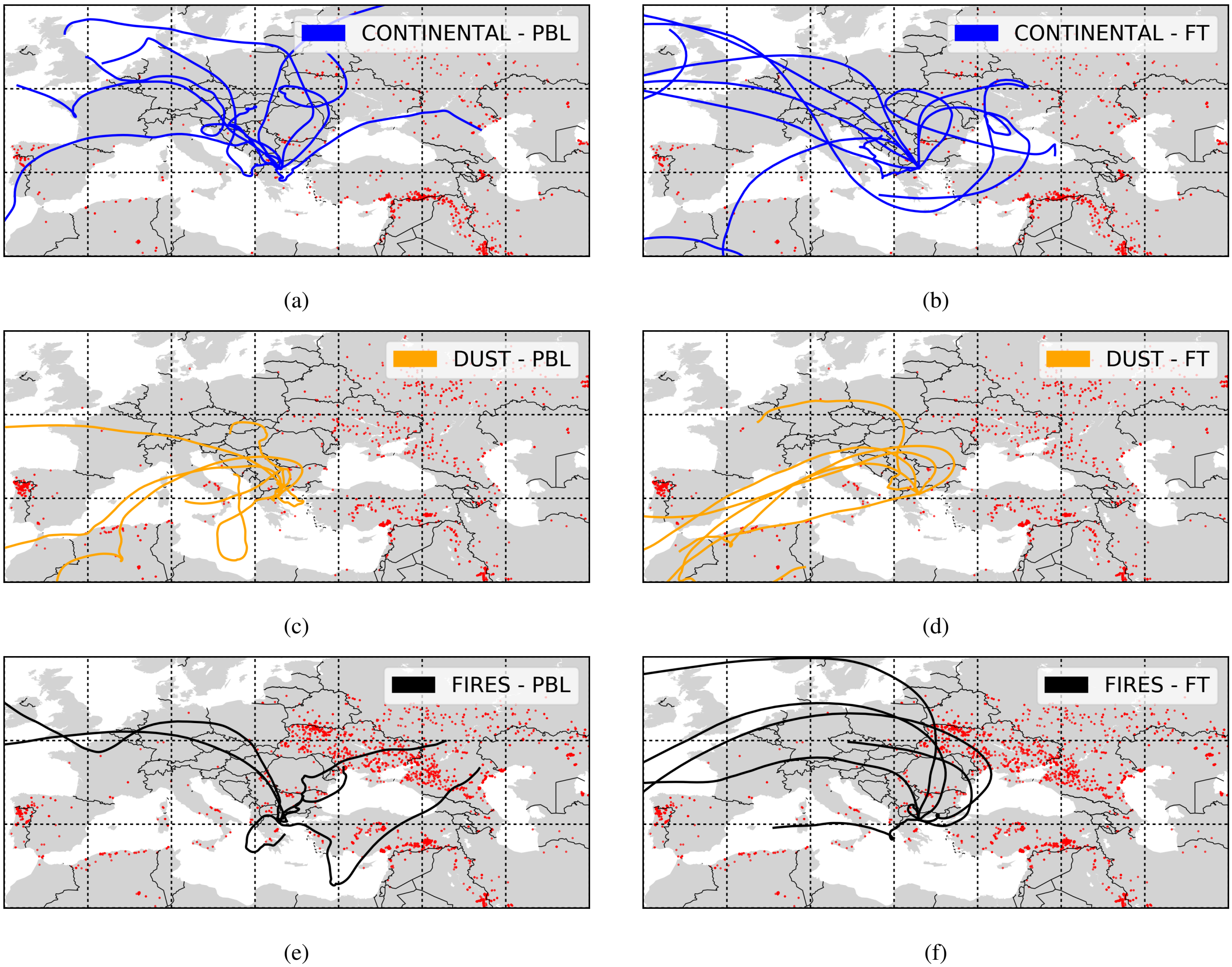


Figure 2. HYSPLIT 6-day backward trajectories and MODIS fire pixels for the continental (a and b), desert dust (c and d) and biomass burning (e and f) cases in the period 2013-2015. The left column includes trajectories that arrive in the boundary layer while the right column includes trajectories arriving in the free troposphere.

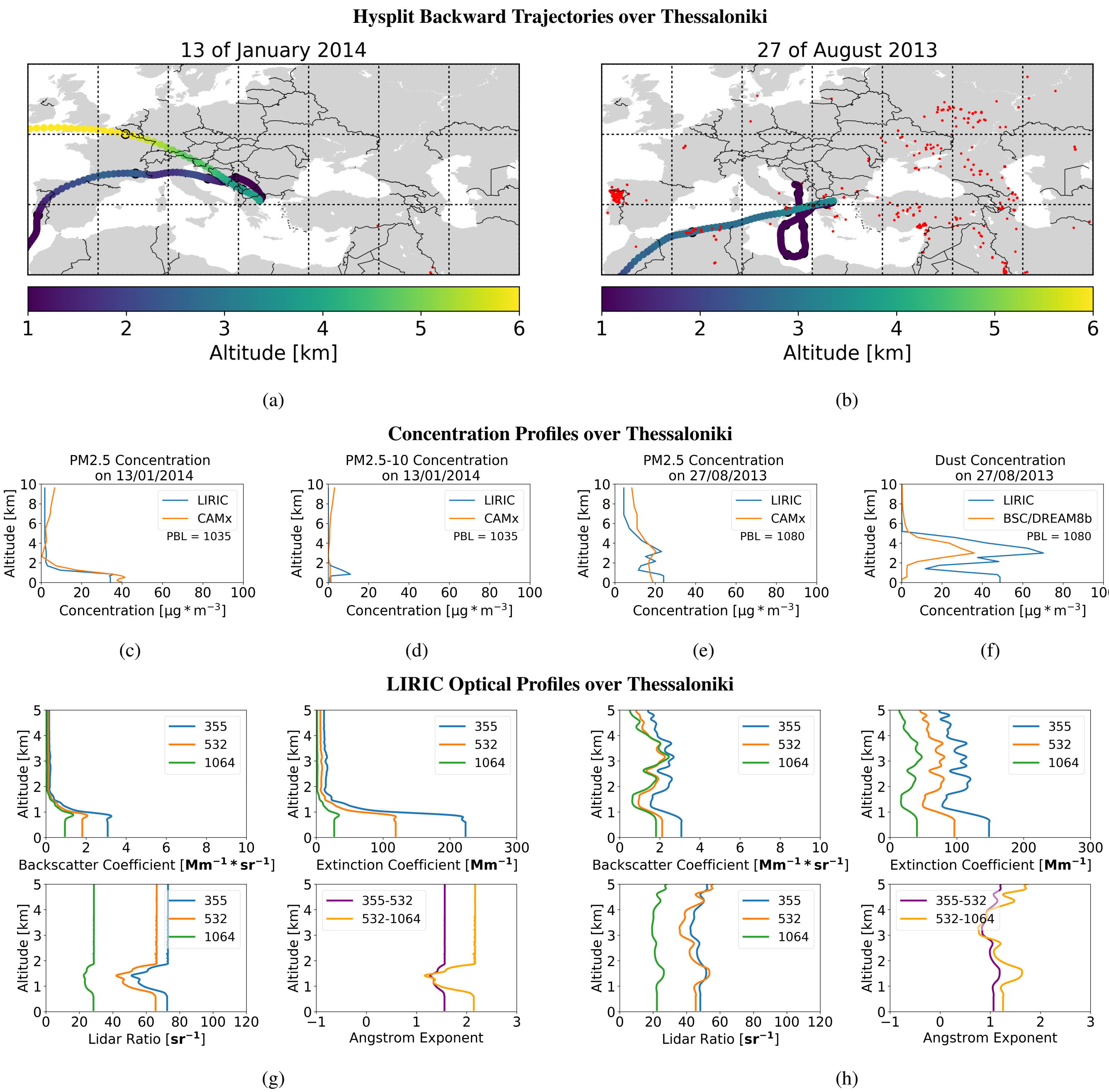


Figure 3. Two sample cases are presented here. The left column corresponds to a continental aerosol case on the 13th of January 2014 while the right column corresponds to a dust aerosol case on the 27 of August 2013. The wind back-trajectories (a and b), the ~~fine~~ PM2.5 and ~~coarse mode~~ PM2.5-10 concentration profiles of liric (c, d, e and f), and the respective optical products (g and h) are include. The big black dots in a and b indicate 24h intervals

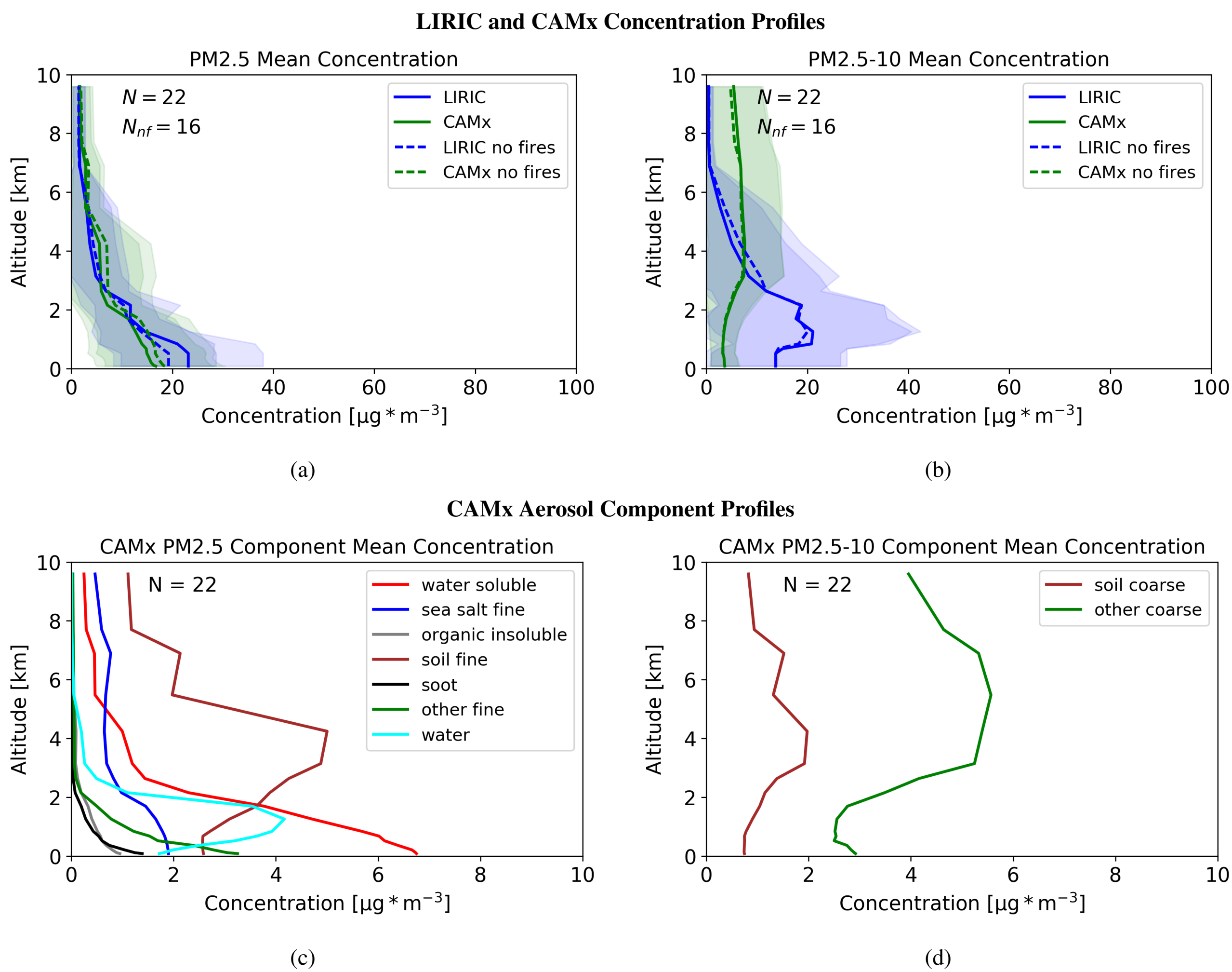


Figure 4. a and b: Comparison of the mean ~~fine mode~~ PM2.5 (a) and ~~coarse mode~~ PM2.5-10 (b) aerosol concentration profiles between LIRIC and CAMx in the period 2013-2015. The shaded regions correspond to one σ of each average profile.

c and d: Presentation of the mean concentration profiles per aerosol component out of which the model's ~~fine mode~~ PM2.5 (c) and ~~coarse mode~~ PM2.5-10 (d) modes are consisted (see table 1).

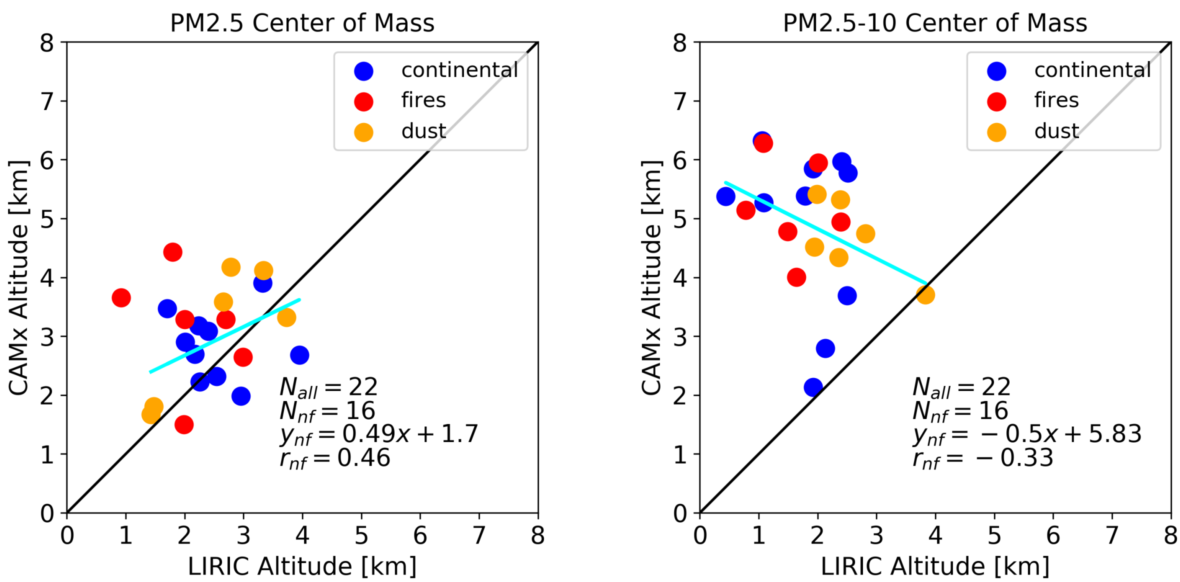


Figure 5. Scatterplots of the fine PM2.5 (left) and coarse mode PM2.5-10 (right) center of mass for LIRIC and CAMx in the period 2013-2015. The biomass burning cases are marked with a red color while the dust cases are marked with orange. The ’nf’ label corresponds to metrics of the category ’no fires’. The least square fit line corresponding to the screened cases (blue and orange) is also included.

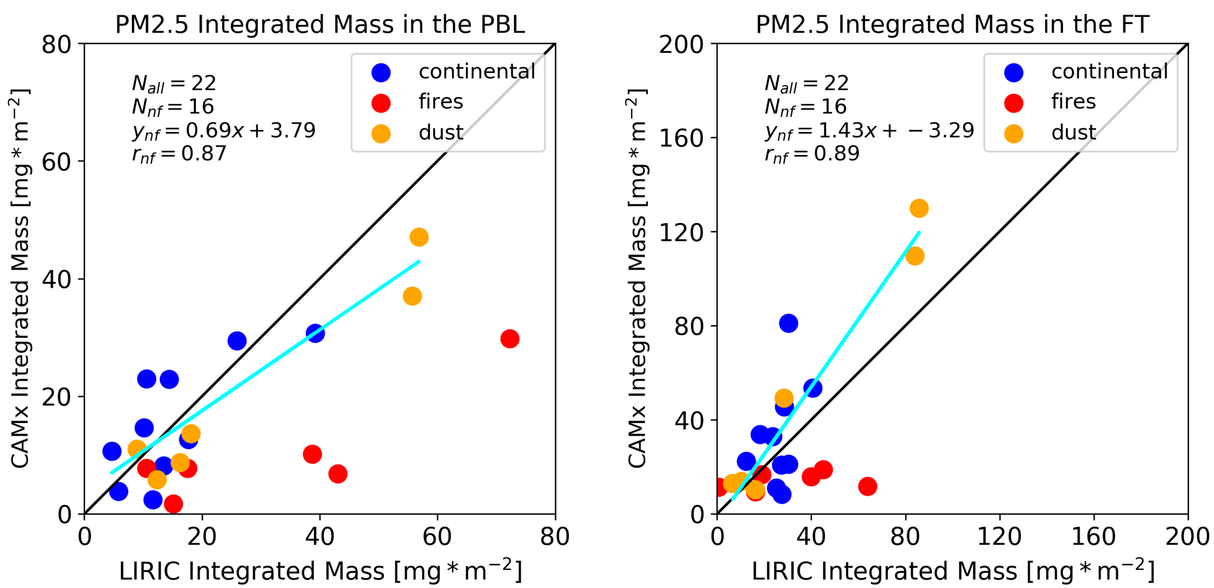


Figure 6. Scatterplots of the ~~fine mode~~ PM2.5 integrated mass in the boundary layer (left) and in the free troposphere (right) for LIRIC and CAMx in the period 2013-2015. The biomass burning cases are marked with a red color while the dust cases are marked with orange. The 'nf' label corresponds to metrics of the category 'no fires'. The least square fit line corresponding to the screened cases (blue and orange) is also included.

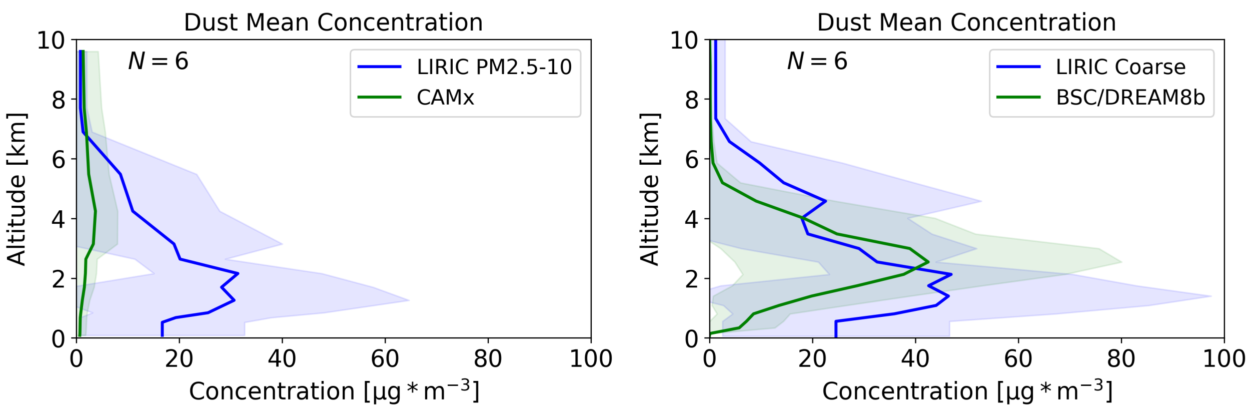
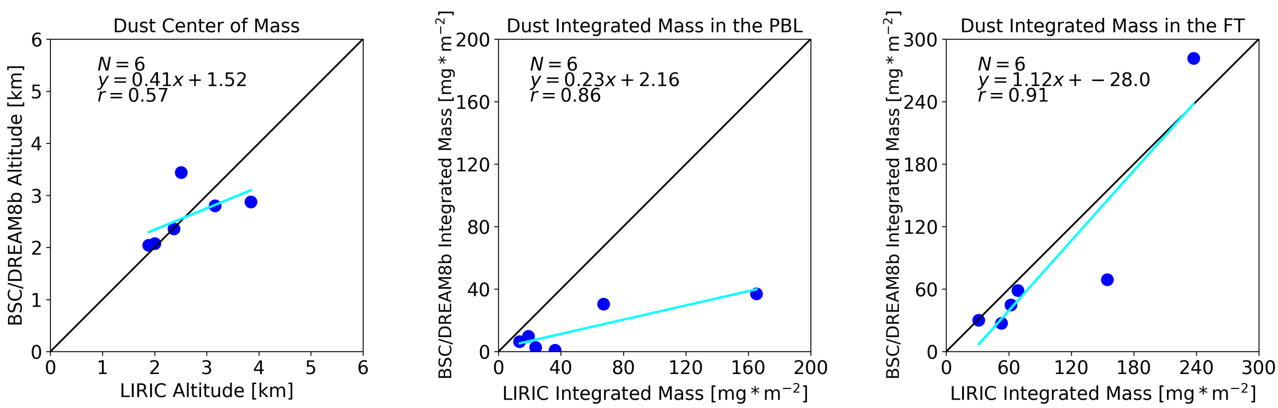


Figure 7. Comparison of CAMX (left) and BSC-DREAM8b (right) mean dust concentration profiles with LIRIC in the period 2013-2015. The mean ~~coarse mode~~ PM2.5-10 profile from LIRIC ~~,~~ is compared with the mean soil ~~coarse~~ PM2.5-10 component profile from CAMx~~and the~~. The mean total dust profile from BSC-DREAM8b ~~are presented~~is compared against the coarse profile of LIRIC. The shaded regions correspond to one σ of each average profile. Differences between the LIRIC profiles are attributed to the interpolation of LIRIC in the resolution of each model.



(a) (b) (c)

Figure 8. a: Scatterplots of the center of mass of the dust cases for BSC-DREAM8b.

b: Scatterplots of the integrated mass of the dust cases for BSC-DREAM8b in the boundary layer.

c: Scatterplots of the integrated mass of the dust cases for BSC-DREAM8b in the free troposphere.

The least square fit line is also included. The correlation coefficient for those cases is provided as well. The figure corresponds to the dust cases of the period 2013-2015.