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 The "fine" and "coarse" terminology has been replaced by "PM2.5" and "PM2.5-10" where it was necessary.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 “between”.

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"