Answer to all referees:

We first would like to thank all referees for their valuable comments, especially about the comparison between models and observations. To answer to all their comments, the authors have had to go through the data analysis again and finally made little changes to improve the comparison. For this reason, the referees will notice that in the revised manuscript the results (figures and tables) of the section about the model comparison have slightly changed. This is due to several reasons:

1. Both models have different time resolutions (NMMB: 3 hours; DREAM: 1 hour). The comparison in the original paper was made with the resolution of each model. However, in order to be comparable, the authors think that the comparison models-observations have to be made at the same time resolution. Because of this we have taken the same time sampling of 3 hours for the comparison of both models. As a result, DREAM mean profiles (figure 13) are slightly different and the standard deviations associated to them are a little larger.
2. The extinction value at a given height, h_i, of the models is the average extinction of the layer comprised between $h_i - \frac{h_i - h_{i-1}}{2}$ and $h_i + \frac{h_{i+1} - h_i}{2}$. In the original manuscript the model extinction value was compared to an interpolated value of the lidar profile at the height h_i. And this was not correct. To correct this, the extinction values of the lidar profiles represented in Figure 13 have now been calculated as the mean values of the original lidar profile (at the lidar original vertical resolution) calculated in the exact same layers of each model. This modification has two effects visible in Figure 13 of the revised manuscript: the lidar profiles are smoother, and the lidar profiles compared to NMMB and DREAM are different (because the heights and resolutions of the models are different).

Extreme event. There is no definition of extreme events in the paper. The extreme nature of the event should be addressed explicitly. How this episode is extreme? For example, what is the frequency of such events over the Iberian Peninsula? Note also that AOD values of 2 and larger are not uncommon over Africa. See for example the papers on the Fennec field campaign.

Answered in lines 91-97

Introduction. The introduction is lengthy. It details many general aspects on dust and its impacts (cloud condensation nuclei (no ice nuclei?), radiative forcing, aircraft operation, health issues, ...) that are not addressed in the paper. I suggest to either shorten these parts or to address these issues for the dust episode under study. The latter option would make the paper much more interesting than it is actually.

We have maintained the objective of the paper but we have not addressed further issues, however we think it is important to mention the different implication of dust on climate.

Origin of dust. The paper does not discuss the origin of dust. This must be done with the objective to better document the episode, by using backtrajectories for example. This would also help to discuss the successes or failures of the forecasts.

The origin of the dust and a better documentation of the episode are now discussed in section 3.1. On one hand the back-trajectories during the period of study are presented, as suggested, and the related discussion introduced in the manuscript (lines 358-373). On the other hand and also as suggested by another review, Fig.1 was modified to include several plots that not only show the geopotential height at 850 hPa, but also the surface wind friction velocity, which is a good indicator of possible dust emissions from deserts. The related discussion is included in the manuscript (lines 294-322).

Sharav cyclone. The low over Morocco looks like a Sharav cyclone. There is quite a numberofpapersdiscussingsuchcyclonesandtheirroleindustemission. References to this literature seems more than welcome for documenting this particular February 2017 dust event in a broader context.

Performance of dust models. The paper shows an assessment of the dust forecasts against lidar measurements, but it is very limited in the possible causes of the model deficiencies. A more thorough discussion on such causes must be provided. Furthermore, the quality of the forecasts should not be limited to the assessment of the vertical profile of dust extinction. It would of a larger interest to discuss the model performance in terms of radiative fluxes (because the importance of aerosol radiative forcing as stated in the introduction), sensitive weather variables (temperature and humidity at 2 m, wind at 10 m) and horizontal winds (because it is a potential cause of model discrepancies as written line 742).

A comment is introduced

Calibration issues. In Fig. 5, the RCS signal presents a large change at 1200 UTC 21 February. So does the signal at 8-km altitude shown on 23 February in Fig. 8. These changes suggest a strong issue on the lidar calibration. Please comment these changes and the data reliability.

Further explanation is given. There are no calibration issues anyway.

Minor comments

Figure 1 shows the mean sea level pressure, with many small-scale features due to orography. In order to describe the synoptic circulation, I suggest to plot the geopotential at 500 hPa, or at 850 hPa.

New plots are introduced in figure 1 (geopotential at 850 hPa and wind friction velocity)

Page14, lines299and300. Figures2band2cdonotshoweasterlyandsoutheasterly winds.

Changed, a correction is introduced

Page 18, line 356. The acronym RCS must be defined here, not afterwards (line 511)

Done

Page 19, line 365 typo on "especial" Page27,line503.

Done

Remove "extraordinary" unless you explain the "extraordinary" character of the event

Done

Figures 5, 8, 10. Please add the days on the time label and use a larger font for all the labels.

Done

Figures 14 and 5 and Table 5. Please specify in the caption for which variable the correlation coefficient and the fractional bias are computed. This remark applies to the text as well.

The correlation coefficient and the fractional bias are computed for the extinction coefficient. It is now said in Section 2.3 about the models description and in the captions of all tables and figures concerned.

Page 38, line "649". Please avoid the adjective "nervousness" for qualifying a meteorological model. Page 39, lines 682-686. Remove the discussion on the

troposphere-stratosphere exchanges as the dust plume is not concerned by this process (or "very unlikely" as you wrote).
Page43,line779. Remove "extraordinary" unless you explain the "extraordinary" character of the event
Done
Page 44, line 800. Remove "perfectly"
Done