Response to reviewer #2 of the manuscript

‘Evaluation of ECMWF IFS-AER (CAMS)
operational forecasts during cycle 41r1 - 46r1 with
calibrated ceilometer profiles over Germany‘

for publication in GMD (MS No.: gmd-2020-308):

We thank the reviewer for her/his willingness to carefully read our manuscript and for the helpful and constructive comments!

General comments:

Following the reviewers recommendations, we shortened the manuscript main body quite a bit from ~43 to ~30 pages (now Introduction: 2.5 p, Methodology: 6 p, Results: 13.5 p, Discussion: 7 p, Summary 1 p, References: 7 p, Appendix 9 p). We removed redundancies and tried to keep close to the main points. We are confident, that this makes the article more attractive also to non-lidar-experts and easier readable now.

We are aware of only very few evaluations of the CAMS model with aerosol profile data and receive many questions to this topic at every open discussion in the community. Thus we think that what we now have in section 2 are central elements necessary to understand and classify the results we show. We carefully worked on minimizing it and think we reached a reasonable balance between complication and sophistication.

We made more use of the appendix section, though, which is now about 9 pages. It disburdens the main text while keeping the information we consider necessary, original and relevant.

Following the recommendations we removed 4 figures: deleted 2 tiny-looking and dispensable figures from the main text, shifted two figures to the appendix and removed two other figures from the appendix.

We removed (former) Table 2 as it was redundant information with Figure 4, and removed part 2 of Table 3 as this information is provided as numbers in Figure 5.

We increased the fonts and annotations of all figures and corrected typos and erroneous axis-titles.

Specific point-by-point:

- Cloud formation: It is beyond the scope of this article to go into details of this process, although we are aware that it is a bit unsatisfactory to mention it without really digging into it. Thus we specify the term ‘water’ cloud formation only for the discussion of the described specific event as this is what we regularly observe at Hohenpeißenberg – formation of stratus clouds in the dust layer near the top of the PBL where ice formation at ambient temperatures is mostly unlikely. According to literature, ice or water nucleation may occur – that’s why we do not specify it where we speak about this process in general. We added it here because SD activation gives rise to a marked bias anomaly, deviating strongly from what we usually observe for IFS-AER during events with non-activated Saharan dust. We have no definite observation of rain suppression yet. Unfortunately we can’t find a really appropriate reference to this (quite novel) topic.

- We tried to disentangle the information in the main text and the appendix by deleting those appendix Figures A1a and A1b which were largely redundant with Figure 3. The we moved the complete cloud formation case study to the appendix as it is quite independent. In the text we refer only briefly to it's most relevant results and implications.

- GALION (Global Aerosol Lidar Observation Network) and (yes!) EARLINET is now mentioned and explained in the revised version

- Section 2.1.1 (mass-to backscatter conversion) is shifted to the appendix. We consider this section important since the forward operator, including observed and approximated meteorological and physical quantities, is a significant source of uncertainty in our analysis to which we refer in the errors discussion.

- Oh yes surely…, Lufft

- See answer to comment on section 2.1.1

- You are right – this was changed

- Yes, the realism of near-ground structure in the observations is an important point: We state repeatedly in the text that we cannot interpret model-observation differences below 300 m a.g. At this height the overlap correction usually is around 50-70% for CHM15k (about factor 2 to 3), which can still reasonably be corrected. In this article the lowest altitude we discuss is thus 400 m above ground. The monthly mean profiles in Fig 3 are averaged over 21 stations and a month such that all random errors of the instrument-specific overlap functions should cancel out. The consistent tendencies found between 400 and 1000 m a.g. are thus considered quite reliable.
- Figure 3 is one of our two central figures which we think is best suited to illustrate the vertical profile bias in the model. It gives a complementary perspective to Fig 2 and we can't really see how this information could be displayed in a more intuitive and condensed way. We decided to add error bars here to the less relevant control run bias profiles because this is the figure where we think the reader gets the best and easiest feeling for the inherent uncertainties.

- In order to tidy up we removed the fire case from the results section and only briefly mention it in the discussion as it is related to the Saharan dust case on 16/17 Oct 2017 and behaves in IFS-AER in a typical way and can thus be easily described as a quite common case.
- The cloud formation case including figures is moved to the appendix. We unburdened the results section through reducing by ~40% while still keeping the most relevant information.

- We thought of removing, shifting-to-appendix or even stronger reducing the mixing layer height discussion, but decided to keep it because of the repeatedly experienced large interest for this topic from the scientific (also non-aerosol) community. Even the fact that operational algorithms at present do not provide reliable results for most of the data is confirmed to be relevant in the discussions we have. Thus we think it a useful and keepworthy result that the mixing layer height from the NWP output is remarkably consistent with a manual analysis.