An intense Saharan dust outbreak was observed with two lidar systems and a sun photometer. The observations are compared to model prediction by ICON-ART model. Arrival time and dust layer heights agreed for the first dust plume, for the second dust plume at higher altitudes the agreement is less good. The backscatter coefficient was overestimated by the model, the AOD underestimated. A lot of work was put into the lidar analysis for one vertical pointing and one slanted (30°) and vertical pointing lidar system. The structure and the language needs improvements.

Comparing a single dust event with the model predictions of a single model is surely a lot of effort, but it is not state of the art anymore. Therefore, major revisions are necessary before publication.

I would consider a single dust event evaluated with different dust transport models more interesting for the community. Doing so, the strengths and weaknesses of the models could be pointed out. At least two or three more dust transport models should be compared.

Or evaluate multiple dust events with the same model to get some statistics when the model predictions are in line with the observations and when not and to look for the reasons. One event evaluated with one model is surely not enough for a publication in 2021.

Already much more complex publications concerning the comparison of ground-based remote sensing (lidar) and dust transport models are present in literature, e.g. for a dust plume across the Atlantic Ocean (e.g., Kanitz et al., 2014), extreme dust events (e.g.,
Solomos et al., 2017), year-long statistics (e.g., Mona et al., 2014), multi-station statistics (e.g., Soupiona et al., 2020) and fine and coarse dust mass concentrations (e.g., Ansmann et al., 2017). It seems that the present study was performed without the knowledge of the progress made in the past decade.

The list of dust transport models mentioned in the introduction is not complete and should be updated. NMMB/BSC-Dust and SKIRON came immediately into my mind, but certainly there are more.

Literature is full of comparisons of models with observations. A more careful literature research is definitely necessary to place your observations in a broader context and to clearly state the novelty of your study.

Major comments:

- The range limitations of KASCAL (Fig. 1) were not discussed. No data were reported above 6 km. Why?
- Section 3, lines 149-180: The backscatter coefficient itself does not tell you, which layer is dust and which not. Throughout the section, you are writing “dust”. At this point, you haven’t shown the depolarization ratio yet to demonstrate, that your measured backscatter coefficient is really dust.
- Line 171: This statement is true for the first dust layer arriving on 7 April 2018. However, the second dust layer arriving in the evening of 9 April 2018 at around 5-6 km height was predicted by the model around 12 hours too early.
- Lines 176-177: How do you calculate the overestimation of the backscatter coefficient? Please provide more detail on how you get to that value.
- Line 211: Why do you chose lidar ratios of 30 and 50 sr? You measured different ones, reported some lines earlier.
- Line 238: Why do you use an AE of 1? You have measurements of the actual AE.
- The discussion about the single scattering albedo (SSA) and its comparison to literature values could be omitted. You use the standard AERONET output and compare it to literature values. We do not gain additional information out of it.
- Please state at some point, that you are referring to linear depolarization ratio (in contrast to the circular depolarization ratio).
- Your conclusions (lines 282-284) are just qualitative. Please consider some more quantification of the model performance. Therefore, you would need more observations for comparison. Or you would need different models for the same dust event to compare the different model outputs and quantify the agreement.
- If the model could not even reproduce the columnar values such as the AOD (lines 291-297), I am not so confident, that the model compares so well as you stated. To reproduce the arrival time, simply the meteorological fields are correctly predicted. However, if the optical properties such as backscatter coefficient, AOD and AE (see Fig. 7) are not agreeing, the model is probably not the best choice for dust predictions. Here, a comparison to various dust transport models would be nice.
Minor comments:

- The English language has to be checked again, especially in the introduction.
- Keep a uniform format for the dates throughout the paper.
- Line 30: You are not studying dust-cloud interactions in the present study.
- Line 42/43: The sentence could not be understood without the knowledge of the location of SAMUM-1 and SAMUM-2. Please provide the locations and distance from dust source.
- Line 72-74: Why do you mention the results from the North Pacific? There is almost no connection to your work.
- Line 244: The sun photometer works only under clear sky conditions as well.
- Lines 251-252: Please check again Freudenthaler et al., 2009. The values reported at 355 nm are lower than 0.33.
- 1: Is there a special reason to show this specific profile? Please indicate the time of the profile with a vertical bar in the time-height plots on the left. Are the heights reported above ground level or above sea level?
- 2: The figure is too small. Details can’t be spotted. With elastic method, you mean the retrieval using the Klett algorithm? Which filter was applied to the plots shown?
- 4: Which are the uncertainties for the lidar and the model output? Please add error bars to the data points. Caption in not complete, model (green squares) is missing.
- 5: Which smoothing was applied to the data?

References:


