Comment on acp-2021-219
Thomas FLAMENT (Referee)

The paper presents a large event of dust transport originating from the Sahara and crossing the tropical Atlantic ocean westward, until it reaches Central America and the Caribbean sea.

This event was large in the amount of dust transported and was reported in newspapers (e.g. https://www.nytimes.com/2020/06/22/science/saharan-dust-plume.html). Such a large plume is an easy target for Aeolus detection of aerosols and is one of the first uses of this High Spectral Resolution Lidar optical properties product. For both these aspects, this paper makes an interesting contribution to the observation of the atmosphere and is relevant to the journal scope.

The structure of the paper is clear. The syntax and grammar are understandable but will require an extensive work to make it easier to read.

The event is large enough that the transported dust is visible without a sophisticated analysis. However, the method used to match CALIPSO and Aeolus data, to estimate the mass flux and the associated errors needs to be described in more details. Many mathematical details are missing to provide a complete understanding of the work. The flow charts give a good overview of the procedure but are not enough.

We need a reference about the L2A processor. I would suggest Flamant et al. (2008, https://doi.org/10.1111/j.1600-0870.2007.00287.x). Alternatively, we, the L2A development team, are currently preparing about the current L2A product which will be submitted soon to the Atmospheric Measurement Techniques journal. It will present up to
date information about the product.

Also, the Aeolus L2A dataset was not widely used yet: more information needs to be provided about the content and overall data quality of the product. Additionally, a discussion of the quality controls applied by the authors would be welcome.

On a technical aspect of data analysis: we recommend to use the "mid bin" product of the SCA. The authors do not mention the existence of two products in the SCA. This is especially important when looking at extinction coefficients, which are more sensitive to noise and better retrieved through this averaged version of the algorithm.

As few papers are available on the subject, this is a technical detail that needs to be mentioned. See also Baars et al. (2021, https://doi.org/10.1029/2020GL092194, “As for the extinction, the SCA mid bin algorithm is to be preferred against the SCA for the lidar ratio analysis.”).

**Detailed line by line comments:**

I. 145-147: It is right to use the SCA rather than the ICA (this is what we recommend as developers of the product).

However, as said above, the SCA provides two sets of extinction and backscatter coefficients. If the extinction is needed, I recommend using the “mid bin” solution. If this was not the case in the current analysis, this might change results significantly. Known defects of the SCA algorithm will be discussed in the future paper we are preparing. In the mean time, I am available to discuss this directly if necessary.

I. 152: “Since the footprints of Aeolus and CALIPSO are not exactly matched, the missing wind data between their tracks have to be filled in using the ERA5 wind field data”. This sentence is very unclear. Does that mean that you use ERA5 winds at the location of the CALIPSO track? This whole “Methodology” section needs to be rewritten with much more detail.

I. 153: “and the measurement uncertainty is on the order of 20%.” Which measurement uncertainty is discussed here? Aeolus estimated errors are often larger than 20 %.

I. 160:
- References are provided regarding the determination of aerosol mass (namely, Müller et
al. 1999) but a discussion about the method is needed. For instance, Müller et al. say the algorithm can cope with errors of ~20% on the lidar data. Is this verified here?

- The inversion method requires information at three different wavelength. What is the procedure to match the CALIPSO and Aeolus profiles? (e.g. advection of CALIPSO profiles towards the Aeolus profile using ERA5 winds?)
As the time and space matching of the two observations cannot be perfect, how is the mismatch propagated into the error estimate?

The combination of the two satellites observation is really interesting but the method really needs to be described precisely enough so that the results can be reproduced.

l. 161: The authors reject data where meteorological data has more than 90% of relative humidity. I suppose this is the method they chose for cloud screening but this needs to be stated explicitly. As said in the general comments, more information on data quality control and selection needs to be provided, e.g. Do you integrate all of the particles in a given cross-section? Could you discuss contamination by particles other than Saharan dust? (clouds, marine aerosols ...)

l. 163: The mass flux is said to be derived by eddy covariance, but it is not clear to me why we would consider turbulent transport and how this would be done. Is the mass flux derived as the integral of m.v for each pixel? Or is it " m'. v' " as stated? If it's actually the second option, how are m' and v' derived?

As a second point, does this requires an interpolation of m and/or v on a common pixel grid? How is this done?

Fig. 4:
- Could you show the lidar profile from which the location of the stars is chosen?

- Using squares triangles and circles separately for each starting point would allow you to pack more information into this figure (e.g. northernmost point is associated with triangles, middle point with squares etc.).

l.192: “deposition” is ambiguous. The HySPLIT data shows a downdraft of the air mass. The dust could also settle while the air mass has no vertical movement.

Fig. 5:
- Why show 4 lidar profiles but only exploit 3 of them? If it’s not necessary, removing the first one would make the figure easier to read. Overall the 3D figures are pretty but contain too much information to be fully readable.

- On panel (a), could you label the date and time of each satellite overpass?

l. 205-206: The origin of these numbers and of the associated error needs to be detailed (cf comment on l. 160).

l. 225: same as l. 161, I suppose this is a method for rejecting cloud contaminated pixels. It would be better if this was stated explicitly.

l. 233-234: how can the dust concentration be highest in the cross-section 3, farthest from the source. I would expect dispersion of aerosols rather than concentration. Could this be justified?

Fig. 7 and Fig. 10: Estimations of mass flux are provided with an error bar but it is nowhere described how these error bars are derived. We need some proper description of this.

Fig. 8: There is too much on this figure. It is hard to read. You need to simplify the presentation.

l. 276-284: It is no obvious from Fig. 11 that the low chlorophyll concentration is due to a lack of iron. More references would be needed to support this claim.

Conclusion: The L2A data was not available publicly at the time the article was submitted. Authors could state that they were allowed to access the data through their participation as a Calibration and Validation team.