The authors use a data set of airborne lidar and in-situ measurements to study the effect of aviation on the optical and microphysical properties of natural cirrus clouds. In an earlier study by some of the same authors, lidar measurements of the particle linear depolarization ratio (PLDR) were used together with backward trajectories to identify cirrus clouds that likely formed in regions of high and low air-traffic density. The present work follows up on the previous study by adding the analysis of coinciding in-situ measurements of the ice crystal size distribution and the ice crystal number concentration (ICNC) related to the two PLDR modes to the investigation of an indirect aviation effect on mid-latitude cirrus clouds. The authors conclude that cirrus clouds that are affected by aviation as indicated by higher values of PLDR also show larger effective ice crystal radii and lower ICNC compared to unperturbed, low-PLDR cirrus.

The authors address an interesting topic that is certainly relevant to the readers of ACP. They have a unique set of airborne observations at their disposal. However, the study itself as well as the presentation of the results need considerable improvement before they can be accepted for publication in ACP. Below is a list of major and minor comments to the authors.

We thank this reviewer for the helpful suggestions to improve our manuscript! The answers to the reviewer's comments will be given in italic and bold.

Major comments

- The text is quite repetitive, often imprecise, and sometimes just confused. Please refer to the minor comments below for details. The authors should review the text carefully. Language editing is urgently needed before publication.

We have carefully revised the text following the suggestions.

- The description of the data set and methodology is rather sloppy:
  - What is the procedure for ensuring that lidar and in-situ instruments have probed the same cloud? How have sections of the flights during ML-CIRRUS been selected for the analysis presented here?

We added the following paragraph to explain, how the flights during ML-Cirrus have been selected for this study:

'However, only eight of the 16 flights were designed in a way, that they provide coordinated lidar and in-situ measurements. The sampling strategy during these flights were as follows: First, the HALO aircraft flew at higher altitudes for sounding the cirrus clouds with lidar (lidar leg). Subsequently, the cirrus clouds were probed by in-situ measurements at several flight altitudes within the cirrus clouds (in-situ leg). […] For our study only these eight flights with coordinated lidar and in-situ measurements are relevant. Information on the flights (including their Mission ID to make it comparable to Urbanek et al. (2018)) are given in Table 1.'

  - Does the data set include all cases of coincident lidar and in-situ measurements? What is the volume of the data set of coincident lidar and in-situ observations? Is it a few minutes or several hours? This information is also missing in the corresponding plots.

We used all coincident lidar and in-situ measurements for this study. And added the following sentence to provide information of the duration of the measurements:

> ***'Typical lidar legs took about 30 min to 50 min; with a typical aircraft speed of 200 m/s that result in an observed cloud dimension of about 360 km to 600 km. The in-situ legs took a minimum of 10 min per constant flight altitude.'***

- Instead of providing information on the water vapour measurements (which are not used at all in this study), the authors could define the backscatter ratio for lidar non-experts or state what typical cirrus PLDR values are.

We removed the information on the water vapor measurements and included information on the backscatter ratio and the PLDR.

- How are regions of high and low air-traffic density defined and how is the connection made to the measurements? You could provide a quick review of the procedure in Urbanek et al. (2018) and clarify that your analysis is the continuation of their work based on the same cases. If this is clear, you wouldn't need to reproduce an already published figure that isn't really necessary here (as is obvious from the fact that it isn't discussed at all). In any case, figures shouldn't be reproduced from an earlier publication.

We followed this suggestion and clarified that this analysis is the continuation of the work by Urbanek et al. (2018). We thus removed the figure that was already published in this former study. Furthermore, we included a short review of the procedure by Urbanek et al. (2018) to define high and low air-traffic density:

'Urbanek et al. (2018) grouped these flights respectively if the cirrus clouds developed in regions with enhanced background aerosols due to aviation or in rather pristine regions. Therefore, they used 24-hr backward trajectories calculated with the trajectory module of CLaMS (McKenna et al., 2002). They used the maximum cloud ice water content to determine the most probable location of the cirrus development and compared that to maps of enhanced background aerosols due to aircraft emissions (Stettler et al., 2013).'

- The section on the in-situ measurements is quite confusing and ends with a statement that comparison of the data sets shows good agreement. While we don't know what that means or how the data of the different instruments have been combined, it seems that the consideration of data from NIXE-CAPS is sufficient for the purpose of this study. Please stick to the necessary data to keep the study simple.

We rearranged the text in the manuscript to emphasize and clarify the use of the different cloud probes. We used mainly the combination of the CAS, CIP and PIP for the data evalution. We also added a sentence that the data of the combination of the CAS, CIP and PIP and the NIXE-CAPS are redundant.

- The analysis of cirrus measurements for different temperatures could already be motivated and outlined in the methodology section. The availability of in-situ measurements of temperature and humidity is first mentioned in line 168. Why are these measurements and the corresponding instruments not listed in the data section?

We now include these measurements already in the data section.

- The authors should revise the presentation of their data and results:
  - Table 1 is a reproduction of parts of Table 3 in Voigt et al. (2017). It would be more useful for the reader if the table was to include information on the inferred parameters for different cases rather than generic mission information. I would like to see, for instance, number of data points, length of measurements, mean PLDR, mean ICNC, mean Deff, etc. For readers of a scientific paper, dates – as used later in the plots – are much more tangible than arbitrary mission IDs.

The reviewer is right, that the Table in the current version does not provide enough information. We included the median PLDR as given in our previous study (Urbanek et al., 2018) and added information on the temperature of the in-situ legs. However, we could not give mean ICNC and mean Deff as we did not look for that for specific clouds but for specific temperatures. We thus added information on the number of data points for the temperature steps in Figure 8.

  - Figure 1 is a reproduction of Urbanek et al. (2018). Why not just state that their cases will be used for closer inspection in this study? In any case, the authors should briefly review the approach in Urbanek et al. (2018) for allocating high and low PLDR cases to regions of high and low air-traffic density, respectively.

We removed Figure 1 and clearly identify that this is a follow-on study. This also includes a brief review of the approach used by Urbanek et al. (2018) (see comments above).

  - It is not clear what the readers should take from the lidar plots in Figures 2 and 5. It would be more straightforward to combine Figures 2c and 2d with Figure 3 (and analogous Figures 5c and d with Figure 6) into key plots that present all frequency distributions. In any case, please choose a plot mode that allows the readers to separate the different cases. Curves without fill would be an obvious choice over overlaid bar plots. Please also add information on the data sets to the plots if you decide against presenting them in a table.

Thanks for this comment, but we want to keep Figures 2 and 5 in the current form. However, we follow the advice and modified the overlay plots.

  - Figure 4 could be omitted. Its content is fully described in a few sentences. Instead, it would be nice to get more information about the cirrus classification scheme either in the discussion of the second case study or already in the methodology section.

We removed Figure 4 but included a brief description of the classification scheme.

  - I suggest to motivate the analysis of clouds at different temperature already in the methodology. Figure 8 could be discussed before Figure 7. In addition, the left panel in Figure 7 seems unnecessary in light of Figure 8 and its discussion. Stating the message of that figure in the text is sufficient. It would be very helpful if median values were also marked in the plots in Figure 8, for instance by vertical lines.

We followed this suggestion and removed the current left panel of Figure 7 and included information on the median values. However, we included a new panel

showing the comparison of the temperature range where in-situ measurements of both cloud-classes is available.

- The data show differences in ICNC and Deff for clouds with different PLDR. The value of this finding would be much increased if the authors where to present the result of a significance test.

Thanks for this advice. We now include the results of a significance test in the supplement.

Minor comments

- line 36: Bräuer et al. (2021a) doesn't seem to be related to effects of biofuel. Please remove.

Done

- line 38: An increase in ICNC is also found in Marjani et al. (2022): https://doi.org/10.1029/2021GL096173.

We included this reference.

- line 40: IN should be INP as introduced in line 275.

Done

- line 53: Redundant. This has just been clarified in the previous paragraph.

Done

- line 55: ML-Cirrus or ML-CIRRUS?

Corrected

- line 57: pristine really means air-traffic free or with low air-traffic density. Please clarify throughout the manuscript.

Done

- line 64: less cirrus formation? Do you mean a reduction in cirrus occurrence?

Yes, we corrected that.

- line 69: energy forcing = radiative forcing?

In their publication they differentiate between radiative forcing and energy forcing. For the contrails they refer to energy forcing.

- line 75: They also found... Not clear what is stated here.

We modified this sentence.

- line 151-156: This introduction to the first case would be more trustworthy if the authors had provided a review of Urbanek et al. (2018) earlier in the manuscript. Please provide link to Urbanek (2018).

We included review of Urbanek et al. (2018) earlier in the revised manuscript.

- line 160-164: Consider putting this into a table.

The median values of the PLDR are now given in Table 1.

- line 164: It is not clear how embedded contrails have been identified.

The embedded contrails have been identified using the CoCiP (Shumann 2012) calculations. We only rely on these calculations and did not include measurements to identify the embedded contrails.

- line 170-173: As raised in the major comments, this statement indicates that data from NIXE-CAPS should be sufficient for the purpose of this study then.

Thanks for pointing out, that is point has not been very clearly explained. The study was conducted using the data from the combination of the CAS, CIPg-UniM and the PIP. The probe combination offers a measurement size range between 3 and 6400 μm. Deff never exceeded a value of 200 µm, therefore only the CAS and CIPg-UniM were the main instruments feeding into the analysis. This combination has been used in various other publications (Righi et al., 2020, Wang et al., 2023). We added the NIXE CAPS data (with the same size range) for one flight where no data from the main-probe-combination was available due to a failure of the CIPg-UniM. The data sets agreed well when compared for selected flights. We clarified this in the manuscript

line 175: Significant digits: give either exact numbers or rounded values, but not a combination of both.

Corrected

- line 183-186: Please revise the description for precision. Is it a cloud or a flow? What's the dimension related to several kilometres?

We corrected this.

- line 199-205: Consider putting this into a table.

The median values of the PLDR are now given in Table 1.

- line 254: This cannot be seen in the current Table 1. It could be, if the authors where to include a table that presents the values of the considered parameters for the different cases.

We revised Table 1.

- line 303: It would be nice to list these open questions as motivation for further research

We included one of the most urgent question.