Reply to referee #1

We thank Referee #1 for reviewing the manuscript and the valuable comments and suggestions which we address below. The responses to the referee comments are given in blue italic letters.

The authors present an analysis of remote sensing retrievals that were collected during several marine cold-air outbreak (MCAO) flights over the Norwegian Sea. Retrievals are based on passive hyper-spectral and polarization measurements in the shortwave spectrum. Translated into a quasi-Lagrangian framework, the retrievals generally show a progressive deepening of the marine boundary layer and an increasing ice fraction with greater distance from the pack ice. The various MCAO strengths covered by the flights show more intense deepening and greater ice fractions for stronger MCAOs. The paper is generally well-written. However, there are a few concerns that the authors should address before the paper is published. I recommend returning the paper for major revisions.

Major concerns

Quality issues because of sea ice – In several instances the authors write that sea ice impacts the retrieved cloud properties, but it's not 100% clear which data points in the figures are affected. I think the authors should mark suspicious values (e.g., using a different color) in all figures. Also, looking at case "20220404" in Fig. 1, there are substantial flight portions above sea ice, but lines in Fig. 5 and 6 are uninterrupted (and without artifacts); I'm not sure at which times and distances these portions are, but it would be helpful to highlight them.

Thank you very much for noting this issue. In fact, only the retrieved cloud fraction and ice fraction are (partly) affected by sea ice and this issue only applies to measurements in the marginal ice zone. Other retrieved quantities and measurements outside the marginal ice zone are not impacted. The cloud fraction and ice fraction both rely on the cloud mask from the measurements of the polarization-resolving cameras, which are sensitive to the visible wavelength range only. Detecting clouds in the presence of sea ice (especially in the marginal ice zone) correctly is a well known issue in passive remote sensing. We applied a sea ice mask to focus on measurements above open ocean only to avoid additional uncertainties due to the influence of the surface and to filter the most uncertain data. Nevertheless, some outliers still remained in the data. The threshold value of 80% was chosen as a compromise between avoiding artifacts and not excluding the important very initial phase of the cloud evolution. Small variations of this threshold value did not significantly affect the results, but added or removed outliers and smaller or larger portions of the measurements of the very initial evolution.
The lines in Fig. 5 and 6 show composites of the measurements as a function of time and distance above open ocean, which were computed using the backward trajectories. These are not directly measured time series. Excluding parts of the flight tracks due to sea ice does therefore not necessarily lead to gaps in the composites. Some of the lines in these figures are interrupted, if the respective parts of the evolution were not sampled by the flight track. The sea ice mask to filter the data worked better in some cases than in others, which is why more or less outliers are visible in the data of the different flights. This can, for example, be due to the much coarser resolution of the sea ice data of 1km (which is the highest available resolution) compared to the 10 to 100m resolution of the specMACS measurements.
We adjusted Fig. 1 to show, which parts of the flight were excluded due to sea ice and which parts were evaluated. In addition, we extended and rewrote the discussion of the sea ice mask in Sect. 2.1. Moreover, we added the exact times and distances of the outliers to the discussion throughout the paper draft to make clearer, which data points are affected. For all changes, please see the Latexdiff.

Retrieval assumptions – The authors retrieve many cloud microphysical properties and it's not 100% clear if retrieval assumptions are important to do so. For example, are these retrieval look-up-table (LUT) based and what were the assumed cloud vertical structures when generating the LUT (e.g., was liquid always assumed to be above frozen condensate)? The authors should clarify these structural assumptions in Section 2.1. Furthermore, the authors use several thresholds (e.g., for cloud fraction, the watershed algorithm, and thermodynamic phase, etc.). The authors should quantify the sensitivity to these thresholds. For example, for cloud cover the modelling community often uses a cloud-optical depth > 2 or 2.5; which value was chosen here and would slightly different values substantially alter the results?

Thank you very much for noting that. We rewrote Section 2.1 and added more details to the different retrievals, their assumptions, thresholds, and uncertainties. Please see the Latexdiff for all changes.
The cloud top height is derived using a stereographic retrieval and therefore not dependent on threshold values, look-up tables, or forward operators. A detailed validation based on model data was performed by Volkmer et al. (2024).
The cloud fraction was calculated using the cloud mask by Pörtge et al. (2023), which depends on the brightness of the observed clouds. Of course, the choice of the threshold value to distinguish between cloudy and cloud-free pixels affects the absolute value of the derived cloud fraction. This is a known issue for cloud masks based on passive remote sensing. Variations of the threshold value shift the derived cloud fraction to smaller or larger values, but measurements closer and further away from the sea ice edge are influenced similarly such that the presented analysis of the temporal and spatial evolution is not affected significantly.
Since the retrieval of the horizontal cloud extent is the only retrieval which is not published in a separate paper, we added a detailed description of the retrieval method to the appendix.
The liquid cloud droplet size is obtained from the cloudbow retrieval. It is based on multi-angle polarization measurements of the cloudbow, which is formed by single scattering on (spherical) liquid cloud droplets. The droplet size is derived by fitting polarized scattering phase functions according to Mie theory to the measurements. The retrieval was validated based on synthetic data in Volkmer et al. (2024) and an additional analysis based on synthetic measurements showed that the influence of cloud ice on the cloudbow retrieval is negligible for ice fractions smaller than 0.8.
The spectral ice index is directly obtained from the data by computing the spectral slope of the radiance measurements and no look-up tables or forward operators are needed. The threshold value of 20 was not applied in the retrieval but is only used to interpret the results. This threshold value was derived from radiative transfer simulations in Ehrlich et al. (2008,2009), which showed that the ice index values split up between liquid water clouds and mixed-phase or ice clouds with a threshold value of 20 distinguishing the two groups.
The polarized optical ice fraction instead is obtained by fitting multi-angle polarization measurements from a forward operator to measurements. The forward operator assumes homogeneously mixed clouds. The deviation of the actual vertical ice fraction profile from this assumptions leads to an additional uncertainty. However, based on passive remote sensing, no information about the vertical ice fraction profile can be obtained such that assumptions have to be made. The uncertainty introduced by this and other assumptions was characterized in detail in Weber et al. (2025b).

Value for the wider community – While cloud-top height and cloud cover are common properties that can be directly used in the modeling community, I'm less sure how to translate the other quantities and think the authors should discuss it. For example, ice index and ice fraction inform on various layers near the cloud-top; how would it be comparable to other measurements (e.g., in-situ cloud probes) or model output? Specifically for model output, would a forward simulator be needed?

A forward operator is actually only needed to compare the ice index to model output. All other variables can usually directly be computed from the model output. In particular, for the ice fraction and effective radius of liquid cloud droplets, retrieved quantities from synthetic measurement data using LES simulations have already been compared to the model values to validate the retrievals in

Volkmer et al. (2024) and Weber et al. (2025b).
The comparison to other measurement data is more difficult. First, a very good collocation of remote sensing and in situ measurements is necessary. For a direct comparison, the in situ measurements additionally would have to be collected at the correct altitude close to cloud top, which is very difficult. However, a more statistical comparison is possible. Pseudo-vertical profiles of microphysical quantities can be derived from the specMACS measurements (see for example our second paper Weber et al. (2025a)). These profiles could be compared to vertical profiles measured with in situ probes. Unfortunately, during HALO-(AC)^3 in situ measurements were only collected at a very limited number of constant flight levels. The effective radius of liquid cloud droplets can be measured with different in situ probes. For the ice fraction, the optical thickness of liquid water and ice has to be computed from measurements of the droplet and ice crystal size and the liquid water content and ice water content, such that an equivalent optical ice fraction can be derived from the in situ measurements. The qualitative ice index, however, cannot directly be compared to in situ measurements. In Weber et al. (2025a) we show a (qualitative) comparison to radar and lidar observations and combine in situ and remote sensing observations to further investigate microphysical properties and processes in the observed mixed-phase clouds. We added more information about how to compare the retrieved quantities to model output to the discussion: "The macrophysical cloud properties, such as the cloud top height, can directly be derived from typical model output and be compared to measurements. In addition, the ice fraction and effective radius of liquid cloud droplets can be computed from model output as described in Volkmer et al. (2024) and Weber et al. (2025b). A forward operator is only needed for a direct comparison of modeled and measured ice indices, as the ice index is a qualitative measure of the cloud thermodynamic phase and computed from measured radiances. … Weber et al. (2025a) also includes a comparison to active remote sensing observations and additionally applies collocated in situ observations."

Minor concerns

ll. 35-38 I'm not sure that this vertical structure applies to MCAO clouds that are rather convective.

Thank you very much for noting that. The cited reference discusses Arctic mixed-phase clouds in general without the specific focus on MCAOs. Therefore, we changed the sentence to "Arctic mixed-phase clouds have a typical vertical structure". However, radar and lidar data as well as passive remote sensing observations during HALO-(AC)^3 actually also indicate a more liquid-dominated layer at the cloud top for the observed MCAO mixed-phase clouds, as discussed in the second part of this work (Weber et al., 2025) and in Schirmacher et al. (2024). In addition, liquid cloud tops during MCAOs were also observed by Geerts et al. (2022).

ll. 40-41 I'm not sure how relevant the WBF process is inside MCAOs but would probably guess that riming is the dominant mixed-phase process.

We added riming as a second mixed-phase processes here.

l. 107 Do both instruments cover the same range?

Yes, the field of view of the spectrometers is completely covered by the polarization-resolving cameras, which have a larger field of view. We added this information.

l. 124 Are the results sensitive to this threshold (also see second major point).

The results are not significantly affected by this threshold. See the discussion above.

Fig. 2: It would be useful to list the case date in the caption or somewhere in the figure.

We added the date in the caption of Fig. 2 to 4 as suggested.

Fig. 3 and 4: Maybe add uncertainty around each line (e.g. from percentiles).

We added uncertainty in terms of standard deviation to Fig. 2 and 4 as suggested.

l. 288 Cloud-top height in "20220401" does not appear that high?

The cloud top height on 20220401 increased from around 200m to about 1700m. During other flights we see an increase until up to 2500m. We changed the sentence to: "The cloud top height in panels (a) and (d) increases with time and distance from the ice edge from a few hundred meters to a maximum of about 2.5km, depending on the research flight".

Fig. 5: The cloud radius of "20220329" seems small. Is the mean value perhaps a poor representative here?

According to the satellite image and the specMACS observations, the cloud size on 20220329 is similar to the other research flights. This is also reflected in Fig. 5c and f, where the cloud radius shows generally values in a similar range as the other flights.

ll. 230-231 Related to the exclusion of high solar zenith angles: Do higher angles also mean less vertical penetration into the cloud? I think it would be good to provide approximate penetration depth for all cases (and all distances and times).

Observations at larger solar zenith angles are not possible, since the solar and viewing geometry do not allow for the observation of the cloudbow angular range within the field of view of the polarization-resolving cameras, such that the ice fraction and cloud droplet size could not be derived. So, we did not exclude any existing data here, but there is no data available. We changed the sentence to: "For later times (after approx. 210min), no measurements of the effective radius and the ice fraction are available because the solar zenith angle during this part of the flight was too large for the cloudbow to be observed inside the field of view of the cameras, such that the retrievals could not be applied."
Moreover, we added information about the penetration depth to Sect. 2.1 (see above). The vertical penetration depth varies depending on the solar zenith angle. The signal location for the retrieved ice fraction was analyzed in detail in Weber et al. (2025b) and varied between about 1 and 2 depending on the solar zenith angle. We also checked the variation of the solar zenith angle during the research flight but did not find a correlation between the solar zenith angle variation and the variation of the cloud properties.

Typos

l. 244 Please check this sentence. Can uncertainty be negative and does the uncertainty have an uncertainty?

We added the information about the uncertainties to Sect. 2.1 and changed the sentences to avoid confusion. The given values indicated the mean and standard deviation of the differences between the simulated and retrieved droplet radii in the model-based evaluation.