Authors' responses to reviewer #2 comments

Title: Hydrometeor classification of quasi-vertical profiles of polarimetric radar measurements using a top-down iterative hierarchical clustering method

Notes to the reviewer:

The authors would like to thank the reviewer for the feedback and constructive comments on our manuscript. Comments have been taken from the provided review and listed below in numbered list format, and are addressed directly afterword in sub-bullets, for example:

1. Reviewer Comments

    Authors' Response

        "citation from the revised manuscript"

As the reviewer comments were copied from the original review, the line numbers there will refer to the reviewed manuscript. Line numbers in our responses refer to the revised manuscript.

Figures that are in the manuscript are referenced just with the number, e.g. Fig. 1.

Reviewer #2:

The paper presents a new approach for unsupervised hydrometeor classification using polarimetric radar data and temperature. The suggested procedure uses a hierarchical clustering methodology to determine a number of data clusters that can be objectively distinguished from the multiparameter observations. These clusters can be associated with certain classes of hydrometeors using either in situ observations or general physical considerations. It is important that the study utilizes the quasi-vertical profiles (QVPs) of polarimetric radar variables which capture the vertical structure of the cloud and its temporal evolution with high vertical resolution and statistical accuracy. An other positive aspect of this investigation is the utilization of the airborne microphysical probes to label the identified multiparameter data clusters to characterize hydrometeors with different microphysical properties. The paper is well written and is worth to be published after several issues with the analysis and interpretation of the results are addressed. Here are my comments and recommendations:

1. The authors do not consider the impact of the measurement errors (biases and statistical errors) on the outcome of their classification. There is little doubt that if the measurements are too noisy then the number of objectively distinguished clusters of data is reduced. The

potential biases of ZDR and KDP due to miscalibration of ZDR, differential attenuation, and backscatter differential phase in the melting layer (ML) are not mentioned in the manuscript. The evidence of such biases is indicated by negative values of mean ZDR and KDP for several classes. I strongly recommend at least some discussion about the effects of the measurements errors on the classification results.

We thank the Reviewer for the comment. The noise and errors in the data might indeed influence the resulting clustering and we added a discussion of these influences on the clustering and labelling of the clusters in Subsection 5.2, p. 14., L.30 - p. 15, L. 11.

"Measurement errors may influence the clustering results. As it was shown by Bringi et al. (1990) noise in the observations has a strong impact on k-mean HCA results. Unfortunately, it is impossible to run the same type of analysis conducted by Bringi et al. (1990) for an unsupervised hierarchical clustering algorithm as the added noise might deliver modified hierarchical structure with another optimal number of clusters and direct comparison to the original set of final clusters would be impossible. Here this issue of noise is partially addressed through our use of QVP. In particular, azimuthal averaging of a QVP reduces the noisiness of the differential phase within the melting layer (Trömel et al. (2013, 2014)) and was recommended in Kumjian et al. (2013) to quantify rather small enhancements of Z_{DR} and K_{DP}.

The mean negative values of Z_{DR} and K_{DP} in some clusters (f_cl1, f_cl3 and f_cl7) might point at potential biases due to miscalibration of Z_{DR}, differential attenuation, or backscatter differential phase in the melting layer. Biases in the data such as miscalibrations will not impact the clustering process but will impact the labelling as it is based on cluster characteristics. A miscalibration of Z_{DR} can also be excluded as we routinely perform calibration of this variable. Z_H and Z_{DR} are corrected for the attenuation in the data pre-processing. Biases caused by backscatter differential phase in the melting layer (Trömel et al., 2014) have not been removed and are evident in the cluster characteristics. The influence of the backscatter differential phase needs further investigation. As discussed, not all clusters can be labelled with absolute confidence solely based on the cluster's characteristics and in situ observations can help to verify these initial suppositions."

We also added a sentence (p. 7, L. 21 - 24) mentioning the attenuation correction of Z and ZDR:

"Here K_{DP} is calculated as the linear gradient of differential phase shift, where the phase shift has been filtered to remove non-meteorological targets (ρ_{HV} > 0.85) and progressively smoothed using decreasing length averaging windows and Z_H and Z_{DR} are corrected for attenuation using a linear correction (Bringi et al. 1990)."

2. I do not agree with microphysical labeling of several identified clusters. For example, the cluster f_cl1 is labeled as upper part of ML. However, it is obvious from Table A1 and Figs. 6 and 8 that the corresponding signature is observed at negative temperature above the ML and is likely associated with heavily aggregated or rimed snow. Its melting often produces the sagging of the ML as demonstrated in Fig. 6.

   The cluster labelling is clarified in the manuscript p.13, L. 25 - 32:

   > " These characteristics immediately indicate that f_cl13 can be labelled as the "bright band" belonging cluster having mixed-phase (MP) particles.
   > The MP cluster (f_cl13 in Fig. 6) is observed to have some sagging areas: between 10:00 (UTC) and 12:20 (UTC), around 16:00 (UTC) and near 18:00 (UTC). Note that f_cl1 is observed above the MP cluster f_cl13 exactly at these time intervals (Fig. 6). This sagging "bright band" signature is often observed where aggregation and riming processes are occurring directly above the melting layer (Kumjian et al., 2016 and Ryzhkov & Zrnic, 2019). This suggests that f_cl1 can be associated with the processes of aggregation or riming and labelled accordingly."

3. Cluster f_cl7 has highest KDP which is a manifestation of the rapid growth of ice via vapor deposition and onset of aggregation in the dendritic growth layer (DGL) centered at -15 °C. However, the authors label f_cl9 as DGL although the corresponding temperature is higher than -10 °C. DGL is not a hydrometeor class but a layer where dendrites or hexagonal plates typically grow. Depending on the height of the top of the cloud, DGL can be characterized either by a combination of high KDP and low ZDR or low KDP and high ZDR (Griffin et al. 2018, JAMC, pp. 31 – 50). This important characterization is completely missed in the manuscript as well as the reference to the very pertinent Griffin et al. (2018) article.

   The description of cluster f_cl7 is modified and a reference to Griffin et al. (2018) is added to the text (p. 14, L. 6-10):

   > "The combination of rather high Z_H (17 dBZ) and high K_{DP} at temperatures around -15 °C indicates a cluster with high particle number concentration of small ice crystals mixed with a small amount of bigger aggregates. This cluster is potentially a manifestation of the rapid growth of ice via vapor deposition and onset of aggregation in the dendritic growth layer (DGL) discussed in the details in Griffin et al. (2018)."

   We agree that DGL is not a hydrometeor class, as well as the melting layer (ML). The referencing to these areas in the data as DGL are deleted in the manuscript (p. 13 - 14, p. 32, and p. 34 - 35).

4. Cluster f_cl12 is labeled as "ambiguous small ice / drizzle". Small ice is very unlikely because it would completely melt at T = 3 °C.

   We appreciate the Reviewer's comment on the assignment of this cluster. The mean 3 °C temperature calculated for the centroid of the cluster would definitely mean small ice's melting, but if we look at f_cl12 in Fig. 7 we observe the variation of the centroids of cluster data subsets belonging to different dates. For some dates the centroid of this cluster shows a temperature values of -2 °C or even -4 °C and small ice is very possible at these temperatures. Even so, we recognise the median values that describe the cluster characteristics are indicative of what is recognised as drizzle, so we have changed this in the text (p.14, L. 11 - 12):

> " Combining the low mean Z_H with low mean Z_{DR} (0.097 dB) and temperature about 3 °C we can assume that f_cl12 can be labelled as small droplets (i.e. drizzle)."

and in Table A1 p.34.

5. The onset of melting is determined by a zero value of the wet bulb temperature rather than regular temperature and I suggest using the wet bulb temperature for classification.

   We agree that wet bulb temperature would suit better the purposes of the study but unfortunately the data were not available due to technical limitations. We are planning to work on it in future and will use wet bulb temperature data for the next, more detailed analysis.

6. For a future studies I would recommend using vertical gradient of Z and the height of the top of the cloud as additional classification variables. Vertical gradient of Z better characterizes the aggregation/ riming process than the absolute value of Z. It can be seen from the results of classification presented in the manuscript that the cluster f_cl11 is correctly recognized as pristine ice with low Z and KDP and high ZDR and is identified during time periods when the height of the cloud top was law. Significant aggregation is unlikely in this situation due to low number concentration of ice particles and small difference in their terminal velocities.

   Thank you for the suggestion, we shall endeavour to explore the use of the vertical gradient of Z in future studies and consider how the clusters relate to cloud top height and/or temperature when considering their future microphysical interpretation.

7. I notice that several literature references mentioned in a body of the manuscript (Kumjian 2012, Hampton 2019, 2020, Murphy 2018) are missing in the reference list.

   We apologize for missing this. It has now been added and the reference list has been checked for accuracy.