We thank Alexei Korolev for his helpful comments. These comments helped to substantially improve the manuscript. Below we give detailed answers to the reviewer's comments that are highlighted in *blue***.**

Review of "In-situ Observation of Riming in Mixed-Phase Clouds using the PHIPS probe" by Waitz et al. Overview: The paper presents an interesting study of in-situ observations of rimed ice. This study is essential in view of large errors in cloud models and numerical weather prediction simulations related to the overproduction of graupel. The authors performed a large amount of work manually analyzing over 5000 ice particle images collected in mixed-phase clouds. The results of particle images classification were correlated against bulk microphysical and state parameters. The obtained results could be used for the validation of cloud microphysical schemes in cloud simulations. Besides the merits of this work, I found that the text of the paper requires improvements focusing on clarity of formulations and accurate statements. In my opinion, this could be easily fixed. My biggest concern is the quality of measurements of humidity and vertical gust velocity used in this work. I strongly recommend contacting the data managers of the ACLOUD, SOCRATES, IMPACTS projects and asking them to redirect you to the researchers responsible for the collection and postprocessing of these measurements for further consultation. In my opinion, the paper undoubtedly deserves publication in ACP. However, keeping in mind the issue with the data quality, I do not have a choice other than recommend major revisions.

Recommendation: The paper can be published in ACP after major revisions and addressing the comments listed below

> *We thank the reviewer for the positive general comment. Below we have addressed the proposed revisions. Most notably, we have increased the clarity of the description of riming in the introduction and reworked the analysis of the correlation of riming state and degree with HCR Doppler velocity and relative humidity based on updated and post-processed data.*

Major comments

1. Page 3, 3rd para: The description of initiation of riming is incomplete. It may also start from the freezing of precipitation size drops, e.g., D>100um. Such frozen drops, due to their relatively large fall velocity are efficiently collect cloud droplets with 10<D<40um. It is also worth mentioning that large drops may work collectors of slower falling ice particles (e.g., ice lollies).

> *The phrasing was adjusted accordingly to include frozen droplets: "Riming can be divided into two (not always easily distinguishable) sub-topics: riming of small ice particles (diameter D =100-1000mu m) in clouds and riming of large (1000<D<5000mu m) precipitating ice, graupel, snow particles or frozen precipitation size droplets that collect smaller cloud droplets or slower falling ice particles (e.g. "ice lollies")."*

2. Page 4: The definition of the riming efficiency is concerning: "The riming efficiency of an ice particle is a function of (i) its collection efficiency and (ii) the number of supercooled droplets, integrated over (iii) the time the ice particle spends in the cloud and during precipitation." I could guess that the authors meant "riming rate" rather than "riming efficiency." However, item (iii) assumes integration of the riming rate over time, which yields the mass of the rimed particle. Therefore, item (iii) cannot be used in the definition of the riming rate. When defining the rate of riming, it would be more strict to consider collection kernel in item (i), and droplet size distribution in item (ii). The entire statement should be reconsidered.

> *The paragraph was reworked completely: "In principle, riming can occur everywhere where ice particles and supercooled droplets coexist. Pflaum and Pruppacher (1979) have defined the*

collection kernel of a collector with radius R and a droplet with radius r that have a relative velocity Δv against each other as $K = E1\ E2\ \pi(r+R)^2\ \Delta v$ where $E1$ is the collision efficiency of the two particles and $E2$ the efficiency that the two particles remain attached to each other. Ice-ice collisions can lead to aggregation, droplet-droplet collisions to coalescence and ice-droplet collisions to riming. For riming, these quantities depend on numerous parameters including temperature (Kneifel and Moisseev 2020), humidity (Khain et al. 1999) habit, size and orientation of the ice particle (Ono 1969; Wang and Ji 2000; Ávila et al. 2009) number and size distribution of the supercooled droplets (Saleeby and Cotton 2008) as well as turbulence and vertical velocity (Herzegh and Hobbs 1980; Garrett and Yuter 2014). The number of rime on an ice particle is hence dependent on all these quantities throughout particle's trajectory in the cloud and during precipitation."

3. Lines 156-159: "…data-set were visually classified into seven habit classes: (i) plate-like particles (single plates, sectored plates, skeleton plates, and side planes), (ii) columnar particles (solid columns, hollow columns, and sheaths), (iii) needles, (iv) frozen droplets, (v) bullet rosettes, (vi) graupel, and (vii) irregular particle. In addition to the habits, the particles were assigned the attributes (i) aggregate, (ii) rimed or (iii) pristine." This definition is directly related to the objective of this paper, and therefore, it is worth adding some extra text elaborating on it. It appears that the attribute "non-rimed" is missing here. It is not clear whether aggregates and pristine ice can be rimed or non-rimed. In other words, are "rimed aggregates" or "rimed pristine ice" subcategories of the attribute "rimed" (ii) ice. Into which category will a rimed aggregate fall? To which habit class will aggregates of one column and one plate belong? Should "aggregate" be a habit class rather than an attribute?

Particles can be assigned multiple attributes, so e.g. aggregates can be classified as rimed, too. "Non-rimed" were all particles that were not classified as rimed. However, as the assigned attributes "aggregate" and "pristine" are not used further in this analysis, we removed the mention to unused attributes to discuss only "rimed" and "unrimed" (as well as epitaxial riming state and riming degree in a second classification step).

4. Lines 167-172: "Particles were classified regarding their surface riming degree (SRD) as (i) unrimed (SRD = 0%, no visible riming), (ii) slightly rimed (SRD < 25%, a few scattered droplets on the particle's surface), (iii) moderately rimed (25% ≤ SRD ≤ 50%, up to half of the particle's surface is covered by droplets), (iv) heavily rimed (50% < SRD ≤100%, most or all of the particle's surface is covered by rime) as well as (v) graupel (SRD>>100%, the whole particle surface is covered by multiple layers of rime, so that the structure of the underlying particle is no longer recognizable)." As in the previous comment, this definition requires additional explanations. What is the tolerance of determining SRD = 0% ? For example, if an ice particle has one or two frozen droplets on its surface, which category would it fall? Or what is the SRD for the column shown in Fig.3c? What is the probability of misidentification of ice particles as "SRD = 0%, no visible riming"? Could you assess such probability? The criterion "SRD>>100%" looks confusing for the two following reasons. First, the definition of surface riming fraction as SRD=(area of rime)/(surface area of the facetted ice particle prior riming) limits max SRD by 100%. The case "SRD>>100%" would be relevant to the definition of the rime fraction as (mass of rime)/(mass of ice particle prior riming). Second, in the frame of your SRD definition, how would rimed ice particles with, e.g., 100%<SRD>100%? It is not clear how was SRD identified for particle images with non-transparent sections formed due to refraction. Any explanations in this regard would be useful.

Only particles without any frozen droplets visible on any of the two stereo-micrographs on the particles surface were classified as SRD=0%. The particle in Fig.3c was classified as slightly rimed. This was clarified in the manuscript.

As the particle is imaged from two different viewing angles (120° apart) and due to PHIPS' high image resolution, the classification can be done with high confidence. However, misclassification due to "hidden" frozen droplets on e.g. the back side of the particle in the "blind spot" of the two cameras as well as misclassification due to human error cannot be ruled out. This misclassification probability cannot be reliably quantified. Despite that, we believe *that this manual classification is still the most accurate way to classify SRD compared to AI/machine learning approaches.*

We agree that using this definition SRD>100% is confusing. This was changed to SRD=100% for graupel.

5. Page 6. Along with the description of the particle probes, could you also provide the description of the humidity sensor, temperature sensor, gust velocity probes, radars employed on the Polar-6, Gulfstream-V and P3. Description of the accuracy of measurements (especially the Doppler velocity) would be useful here.

The sensors that performed the atmospheric state measurements are now introduced in Sec.2.

6. Page 8, 1st para: Could you explain why did you limit your consideration by the cases with T>-17C? Could you also indicate the ranges of LWC (min and max values) and Doppler velocity? What is the averaging time of LWC, T, VDopl, etc.? Based on the title of the paper, I believe that all measurements were performed when LWC> LWCmin. It would be relevant to indicate it in this section.

The threshold T>-17°C was chosen as almost no rimed particles were sampled at colder conditions (see Fig.2a).

The averaging time was 1s, this was clarified in the manuscript: "For each PHIPS particle, the corresponding temperature, humidity and velocity data as well as LWC were determined as the average over t=t_s \pm 0.5s around the time of acquisition t_s where each PHIPS particle was sampled." (see added paragraph in Sec.2 as mentioned in the previous comment).

CDP LWC ranged from 0 g/m3 to 0.5 g/m3 and vertical HCR Doppler velocity from -4 m/s to +2m/s. This information was also added to the manuscript. As the analysis presented in this paper purely rely on single particle measurements from PHIPS, no LWCmin threshold was set.

7. Lines 184-185: "The riming rate is a function of the relative flux of available droplets and hence droplet number concentration and relative velocity with respect to the ice particle." This is a misleading statement. For example, following this statement, the riming rate of an ice particle will be higher for the cloud consisting of 1um droplets compared to the cloud with 10um droplets, assuming that both clouds have the same LWC. In fact, the 1um droplet will follow the airstream and flow around the ice particle without accretion, whereas 10um droplets will have a high rate of impact with the ice particle due to its larger mass. Therefore, the above statement should consider droplet size distribution instead of the droplet number concentration.

The definition of riming rate was overworked to include droplet size distribution (see answer to major comment #2).

The averaging of the Doppler velocity (VDopl) over the vertical column is concerning. The vertical Doppler velocity may vary in the vertical direction. Therefore, the averaged along the vertical direction Doppler velocity may be biased compared to that on the flight level. There are many documented examples of the variability of VDopl in vertical direction available from the literature. Could you please comment on the vertical variability of the measured Doppler velocity? It would be useful to see some examples.

The analysis of the correlation of riming state and Doppler velocity was reanalysed based on feedback of the HCR team. According to their suggestion, the cases were first divided into stratiform and convective clouds. The Doppler velocity was then averaged over the gate closest to the radar is used to represent the vertical velocity of each particle, instead of using the whole column average as done in the first version of the manuscript.

8. Humidity measurements. The results of the RH measurements presented in Figs.9, S4f, S6f, S7f are questionable. As shown in these diagrams and stated on page 17, the average value of RHice is centered at ~100%. However, in mixed-phase clouds, water vapor pressure is close to saturation over liquid water, i.e., RHliq=100%. This had been shown both theoretically (Korolev and Mazin, JAS, 2003) and experimentally (Korolev and Isaac, JAS, 2006). After leveling the flight at ~4300m (~12:42:30) the

measurements were performed at T=~-5C to -7C. At that temperature saturation water vapor pressure over liquid water would be equivalent to RHice~105% to 107%. However, as shown in Fig.9 between ~12:44:30 and 12:49 relative humidity over ice was always <5-7% within mixed-phase cloud segments. On the other hand, at ~12:42:30 and 12:44 RHice peaked up to ~150% and 125%, respectively. Such relative humidity over ice would be equivalent to supersaturation over liquid of ~40% and 18%, respectively. These are unrealistically high supersaturation over liquid water, which do not occur in the free atmosphere. Such behavior of RHice is suggestive that the humidity measurements had a major issue.

The RH analysis was reworked using recently published RH data using Vertical-Cavity Surface-Emitting Laser (VCSEL) hygrometer measurements for SOCRATES. The calibration was performed at water saturation by Minghui Diao. In the updated SOCRATES humidity data set there are no more unrealistically high RH_liq values, most being between 95% to 105%, peaking at 110%.

For ACLOUD, the measured RH_liq values were centred around 80% while in cloud, peaking at around 85%, indicating that the humidity had a major issue which could not be resolved. Hence, the data were omitted and the RH analysis is thus only based on SOCRATES data. The RH data for the IMPACTS case study was updated using a quality-controlled version of the data set which resolved the issue with the unrealisticially high RH values in Fig. 9.

9. In situ vertical velocity measurements. As it is seen from Fig.9 (4th diagram from the top) that during descent, the vertical gust velocity (Vz) changed from -15m/s (at the beginning of the diagram) to 0m/s (12:42:30), when the flight was leveled. Such behavior is highly suspicious. It looks that the vertical component of the aircraft speed was not subtracted from the gust velocity measurements. After leveling the aircraft (>12:42:30) the measurements of Vz look reasonable. I am wondering if the inclusion of the time segments with the faulty Vz measurements resulted in a big difference between the fraction rimed ice versus the Doppler velocity (Figs.S4d,S6d,S7d) and versus ambient vertical velocity (Figs.S4h,S6h,S7h)?

The case study figure was updated using a recently published quality-controlled version of the data (Lee Thronhill., https://cmr.earthdata.nasa.gov/search/concepts/C1995869822-GHRC_DAAC.html, https://ghrc.nsstc.nasa.gov/pub/fieldCampaigns/impacts/TAMMS/doc/tammsimpacts_dataset.pdf) but unfortunately, in the period of the case study, all data was rejected. Please note that the IMPACTS data was only used for this case study, the correlation plots (Figs S4, S6, S7) are only based on ACLOUD and SOCRATES data which are thus unaffected by this issue.

10. Epitaxial growth: Line 400: "It is further possible that older rime grows on the expense of recently accreted droplets that partly evaporate due to latent heat during the freezing process." It is unlikely that droplets freezing on the surface of the ice particle may have any significant contribution to the epitaxial growth of the pre-existing rime. During freezing, the temperature of the droplet may temporarily increase to 0C and generate within its vicinity (<2D) a region of high supersaturation, which may potentially enhance a diffusional growth of the rime within its direct neighborhood. The endurance of the enhanced supersaturation is limited by the droplet freezing time, which will be decreased by the heat transfer into the ice crystal. For a 20um diameter droplet, the freezing time is estimated to be less than 50ms. Observation of sparse or single rime regrown into facetted crystals (e.g., Fig.9 (Slightly Epitaxial Riming), Fig.S8(CTA1)) on the surface of collector ice particles does not support the proposed hypothesis due to the absence of freezing droplets around, which may source water vapor for the pre-existing rime. The growth of the rime can be simply explained by the deposition of water vapor available from the supersaturated over ice environment, which is always available in mixed-phase clouds.

The corresponding sentence was removed.

Minor comments

1. Lines 25-26: "Mixed-phase clouds, …, play a major role in the life cycle of clouds…" This sentence sounds confusing. It is worth rewording it. You may consider mentioning e.g., the hydrological cycle.

The phrasing was adjusted accordingly.

2. Line 32: "main growth modes" is worth replacing by "main ice growth modes".

The phrasing was adjusted accordingly.

3. Line 33-35: "Riming can be divided into two (not always easily distinguishable) sub-topics: riming of small ice particles (diameter D = 100 - 1000 μm) in clouds and riming of large (1000 <D<5000 μm) precipitating ice, graupel and snow particles." This statement is disconnected from the rest of the paragraph. A reader would anticipate the following elaboration of this statement and explanation of division in two subranges. What is the relevance of this statement to the objective of the paper? It is not used in the following text.

It was emphasized that whereas most recent publications focus on the latter aspect (riming of large precipitating particles), in this study, we focus on riming of smaller ice particles in clouds. This is also discussed in more detail later in line 67-80.

4. Line 40: "…until gravitational settling becomes efficient." This statement is unnecessary here. Droplets will freeze on the surface of ice particle after enhancement of their gravitational settling as well.

The phrasing was adjusted accordingly.

5. Line 71: "…discriminate between rimed and irregular particles". Consider replacement by "…discriminate between rimed and non-rimed particles". Note rimed particles fall in the category of irregular ice particles.

The phrasing was adjusted accordingly.

6. Line 159: "In addition to the habits, the particles were assigned the attributes (i) aggregate, (ii) rimed or (iii) pristine." This definition is directly related to the objective of this paper, and therefore, it is worth adding time text elaborating on it. It appears that the category "non-rimed" is missing here. It is not clear whether aggregates and pristine ice can be rimed or non-rimed.

"Non-rimed" were all particles that were not classified as rimed. This was clarified in the manuscript (see also answer to major comment #3).

7. Line 186-188: "Further, it is dependent on the collision probability (and hence the cross sections of ice particles and droplets) as well as on the collection efficiency, i.e. the probability that a colliding droplet sticks as rime." This statement has to be modified. Note that collision efficiency is a product of collection efficiency and coalescence efficiency. When talking about the dependence on "the cross-sections of ice particles and droplets", you should also mention ice particle density and its orientation. In general, it is simpler to talk about the collection kernel rather than the parameters affecting it.

The definition of riming rate was overworked to include the collection kernel and coalescence and moved to the introduction (see answer to major comment #2).

8. Line 260, 278, 327: "coder temperatures" replace by "lower temperatures" (jargon: temperature cannot be cold or warm).

The phrasing was adjusted accordingly.

9. Line 340: "warmer temperatures" replace by "higher temperatures".

The phrasing was adjusted accordingly.

10. Line 263: "orientation within the cloud" replace by "actual orientation with respect to horizon"

The phrasing was adjusted accordingly.

11. Line 323: "…investigating the orientation of the freezing of rimed droplets…" change to "investigating the orientation of crystallographic axes of the freezing of rimed droplets.

The phrasing was adjusted accordingly.

12. Line 373: "shown in Fig. 9b". Labeling "a" and "b" are missing in Fig.9.

Labels were added.

13. Line 374: "were not classified since they were identified as potential shattering fragments smaller than D= 100 µm." What was the criterion for identification of shattering artifacts?

Shattering artifacts can be identified from the PHIPS stereo images that have a field of view of approx. 2.19 mm x 1.65 mm by looking for satellite particles. It is acknowledged that shattering fragments do not always appear as "satellites" but can be found as single fragments within the image frame. Such individual shattering fragments can be typically identified as having sharp edges and a shape that does not appear to resemble that of a typical vapor grown crystal (i.e. a lack of hexagonal symmetry of the crystal facets). If such particles were identified during the manual image inspection, they were also categorized as shattering cases. This discussion was also added to the manuscript.

We thank the anonymous reviewer for his/her helpful comments. These comments helped to substantially improve the manuscript. Below we give detailed answers to the reviewer's comments that are highlighted in *blue*.

General comments to the manuscript acp-2020-833

The study titled "In-situ Observation of Riming in Mixed-Phase Clouds using the PHIPS probe" by F. Waitz et al. presents airborne in-situ observations of cloud particles with the PHIPS acquired at three field experiments ACLOUD (Arctic), SOCRATES (Southern Ocean), and IMPACTS (East Coast of USA). The riming state of the ice particles was identified and correlated with temperature, ice particle size, liquid water content, relative cloud height, supersaturation with respect to ice, cloud droplet size, airborne radar mean Doppler velocity and ambient vertical air velocity. Additionally, for a case study of the IMPACTS campaign, the evolution of epitaxial riming (small, faceted rime oriented to the crystalline axis of the host particle) is described in detail.

Without doubt, in-situ observations of riming in mixed-phase clouds under different conditions (different ranges of temperature, liquid water content etc.) as done in this study are valuable for validation of remote-sensing microphysical retrievals as well as for assessing the performance of cloud simulations in different situations. However, several main issues need to be addressed before publication.

> *We thank the reviewer for the positive general comment. Below we have addressed the proposed revisions. Most notably, we have increased the clarity of the description of riming in the introduction and reworked the analysis of the correlation of riming state and degree with HCR Doppler velocity and relative humidity based on updated and post-processed data.*

Recommendation: I would suggest the manuscript to be published after major revisions. The authors should address the following points:

Major comments:

1. The facts and previous studies presented in the paragraphs of the introduction section seem not always well-structured. Some terms are used without explaining what they mean (e.g. l.66 "riming state" and l.76 "riming degree").

> *More explanation was added: "Those methods proved to be fit to determine the **riming state (i.e. whether a particle is rimed or unrimed)** of large, precipitating snow and graupel particles. […] Most models account for the **riming degree (i.e. what fraction of a crystal's surface is covered by rime)** only in the sense of a subtype for hydrometeors (e.g. cloud ice, graupel, snow)".*

2. The way data was grouped and analyzed and plotted needs to be explained more clearly. Specifically:

3. a) l.159: Why is "aggregate" used as attribute to the ice particles and not as particle habit? The term aggregate is introduced here but not used subsequently. Does that mean aggregates were not observed with the PHIPS?

> *In the manual habit classification, we use the attribute "aggregate" together with a habit class. Depending on the cloud situation, information about aggregation can be of value, for example in outflows to distinguish between single habits and aggregated habits, like plate-aggregates. In the present study, however, aggregation is not of importance and not discussed. Therefore, to avoid confusion, we removed the mention to unused attributes.*

4. b) l.167-178: In this second classification step different attributes are introduced (sublimated, one-sided riming, epitaxial riming). How do you assess if a particle is sublimated? How do assess one-sided riming for more opaque particles?

As sublimated particles are not discussed further subsequently, the mention was discarded to avoid any confusion for the reader (similar to the attribute "aggregate", see previous comment). As the particle is imaged from two different viewing angles (120° apart), whether or not a particle is one-sided can also be assessed for opaque particles (see examples in Fig.4). A corresponding remark was added to the text.

5. d) l. 177: You mention the total number of PHIPS particles analyzed for ACLOUD and SOCRATES. To put those numbers more into context of the field experiments, it'd be helpful if you added how many minutes of flight time this refers to.

In total, SOCRATES and ACLOUD consisted of about 330 flight hours. A remark was added to the text.

6. e) The absolute number of particles in Fig2a and Fig S4a are different for the warmest temperature bin even though the caption says that the same data is used. – why?

The temperature plot (Fig. S4a) was using incomplete data. This was corrected in the SI.

7. f) l.201-202: According to Fig 2a, the high fraction of rimed particles of approx. 50% does not happen in a wide T-range of -10°C to 0°C as stated but rather at T of -7 to -8°C. Please rephrase. State how many particles were analyzed around -17°C (instead of "low number", l.204).

The phrasing was adjusted: "Most riming was observed in a temperature range between -10°C<T< 0°C with the maximum around T=-7°C where up to 55% of all ice particles were rimed. The high riming fraction around -17°C is due to a very high rimed fraction during a single cloud segment of RF09 of SOCRATES. It is based on a low number of total particles (n=213) and is therefore not assumed to be a generalizable feature."

8. f) Fig. S1, S2, S3: It has not become clear to me why for the temperature dependent frequency of occurrence distribution of different ice particle habits the data is divided into ACLOUD (S1) and SOCRATES (S2) but then combined for both campaigns in the histogram plot (S3) of the percentages of rimed particles per ice particle habit. – Please explain.

To avoid confusion, the dependent frequency of occurrence distribution of different ice particle habits was combined for SOCRATES and ACLOUD.

9. The section 239-251 on correlation of riming with mean radar Doppler velocity needs revision. All of the experimental studies listed in the introduction that employ vertically-pointing radar Doppler velocity measurements are using ground-based Doppler radar observations. The Mosimann approach as well as Kneifel and Moisseev, 2020 clearly state that their method of using absolute values of mean Doppler velocity to infer riming is only possible if vertical air motion (up- and downdrafts) can be neglected and the mean Doppler velocity thus is close to the ice particle terminal fall velocity. In the present study however, airborne radar Doppler velocity measurements are used instead which are much more complicated to interpret due to the aircraft motion and the two possible radar-pointing scenarios. Depending on where the aircraft was flying with respect to the cloud, sometimes the HIAPER cloud radar (HCR) was pointed upward or downward. According to the description, the radar pointed to zenith (up) when flying beneath clouds or doing ascents in boundary layer clouds. It pointed to nadir (down) "at other times". For the two different configurations, the Doppler velocities towards the radar should thus have different sign conventions OR different interpretations, correct? Please mention if this paid attention to/adjusted in the HCR data before it is used here? Depending on the radar-pointing scenario and the location of the aircraft within the clouds different parts of the clouds are observed with the radar. If for example the HIAPER did an ascend in a boundary layer cloud it pointed upward toward cloud top whereas if it descended it pointed downward towards cloud base. In other words, in the zenith-pointing scenario the radar observes the "history" of the particles that the PHIPS measures

in-situ whereas in the nadir-configuration, the radar shows the "future" of the PHIPS-observed particles. In both cases different parts of the cloud with potentially different microphysical composition and different dynamics (up-/downdrafts, turbulence) are observed by the radar leading to different mean Doppler velocities that are here correlated with riming. This has to be entangled and so simple averages of mean Doppler velocities over the whole vertical column (up-or downward from the in-situ observation) don't seem to be sufficient. Also, the reader wonders about the vertical extent of the clouds probed.

The analysis of the correlation of riming state and Doppler velocity was reanalysed with the recently updated HCR data (version 3.0, https://doi.org/10.5065/D68914PH) and with consultation from the HCR team. Based on feedback of the HCR team our analysis was divided into stratiform and convective cloud segments. The Doppler velocity is no longer averaged over the whole vertical column, instead the vertical velocity of the gate closest to the radar is used to represent the vertical velocity of each particle.

More details on the HCR should be mentioned as well: radar frequency? Temporal resolution of data?

The HCR is now introduced in Sec.2 and frequency and temporal resolution are mentioned.

Also, in the current interpretation of the Fig2d), the mean Doppler velocity seems mostly be explained by dynamics and its influence on residence time of the particles and not the terminal fall velocity of the hydrometeors itself. In that context, please clarify l.249-251.

A statement was added clarifying that the Doppler velocity is the sum of vertical air motion and the fall velocity of (precipitating) particles.

As a minor additional comment, L.242 states that positive Doppler velocity refer to updrafts which should be replaced with "upward direction" as it is again the superposition of vertical air motion (up- and downdrafts) and reflectivity-weighted particle terminal fall velocity.

The phrasing was adjusted accordingly: "The sign was adjusted based on HCR orientation so that negative velocity always corresponds to downward direction, positive to upward direction."

Also, in the description of Fig. 2 it is not very obvious that Fig2d) only refers to the SOCRATES data set (and not ACLOUD+SOCRATES as stated in l.206). This restriction to SOCRATES only becomes clear on l.244-245. The same is true for Fig 6c) and Fig 8c). It should be mentioned more clearly that these plots are based on SOCRATES HCR data only. The description of Fig 6c and 8c should also be revised.

Corresponding remarks were added.

l.252-260: Explain how the vertical air velocity is measured. Turbulence and vertical air motion of large magnitude (strong up- and downdrafts) seem to be not correctly disentangled.

The meteorological sensors are now introduced in Sec.2. It is correct that the ambient vertical velocity measured at the aircraft is the combination of small-scale turbulence and large-scale vertical motion. These parts cannot be easily disentangled. A corresponding note was added to the text.

Minor comments:

-l. 36: replace “following” with “followed by”

The phrasing was adjusted accordingly.

- l.90: “scale” of riming refers to “degree” of riming?

The phrasing was adjusted accordingly.

-l.180: Do you plan to do automated image processing for the large amount of PHIPS images for the entire IMPACTS data set?

Currently, PHIPS image processing exists only in terms of geometrical properties of the ice particles (maximum dimension, aspect ratio, …). This data is available in the HYDRO database referenced in the paper. An automated habit and riming classification algorithm of the PHIPS images is planned for the future but does not yet exist.

-l.160: “distribution” should be more precise: “temperature dependent frequency of occurrence distribution”

The phrasing was adjusted accordingly.

l.161: In Fig S3, the percentage of rimed particles per habit (here: “riming fraction”) is subdivided into “normal”, “epitaxial” which should be mentioned here.

The phrasing was adjusted accordingly.

-l.183: What do you mean by “number of rime”? – maybe “amount”?

“Number of rime” refers to the number of droplets that are rimed on an ice particle (whereas “amount” could refer to the mass). This was clarified in the text.

-l.188-189: ice particle trajectory is also a function of ice particle density and terminal fall velocity. Regarding vertical air motion, “updrafts” are explicitly mentioned but “downdrafts” are also relevant as they decrease the residence time of ice particles in the cloud. Please rephrase.

The phrasing was adjusted accordingly to also mention downdrafts.

-l.194: “in-situ” aircraft data instead of “experimental” data

The phrasing was adjusted accordingly.

-l.195: Subsequently, you also correlate the relative occurrence of rimed and unrimed particles with LWC, ice particle size etc. which are microphysical parameters (not meteorological parameters). Please rephrase accordingly, also in the abstract and elsewhere.

The phrasing was adjusted accordingly throughout the whole manuscript.

-l.196: replace “were rimed” with “experienced riming”

The phrasing was adjusted accordingly.

-l.199: Add an introductory sentence to Fig.2 and move the sentence regarding the fit parameters (l.205) behind it.

The phrasing was adjusted accordingly.

-l.209: acronym HCR needs to be introduced

The HCR is now introduced in Sec.2.

-Table 1: It is not mentioned that the linear fit for Fig2c) (Riming percentage vs LWC) was made with a logarithmic y-axis

The linear fit and all other fits in the other figures were applied to the riming fraction (percentage, right y-axis) which is not logarithmic. This was clarified in the text to avoid confusion.

-l.219: replace “above that” with “For larger ice particles”

The phrasing was adjusted accordingly.

-l.227,230: “rimed droplets” sounds like as if the droplets are rimed which is misleading. Please rephrase. Also it is unclear if the two droplet size examples of 20mum “and” 50mum are representative and show the span of droplet sizes that lead to riming. Please clarify.

Renamed “rimed droplets” to “rime droplets” throughout the manuscript to avoid confusion. The two exemplary particles with rime droplets of 20um and 50um were amongst the particles with the smallest and largest rime droplets based on visual inspection. However, there exists no automated method to determine the size of the rime droplets based on the PHIPS images. Since the estimated size based on exemplary images is in agreement with the cited previous studies, this is not investigated further in this work. This is also clarified in the manuscript.

-l.246: add “clouds” after “flying beneath”

The phrasing was adjusted accordingly.

-l.259: replace “one-sided” with rimed on one side

The phrasing was adjusted accordingly.

l.268: the subsection “riming degree” is introduced here. Consider using the title in l.182 as subsection a) for the first part of the analysis and find a Section title.

The first part of Section 3 (Riming Fraction) was also grouped into a subsection.

-l.281: This sentence seems to contradict the previous statements in line 278-280. Please rephrase.

The sentence was removed.

-l.360: It looks like a subsection for Section 4 is introduced here. The previous analysis in Section 4 should then also be grouped into a subsection.

The first part of Section 4 (correlation with ambient parameters) was also grouped into a subsection.

Comments on Figures in Supplement:

S1, S2,S3: It has not become clear to me why for the temperature dependent frequency of occurrence distribution of different ice particle habits the data is divided into ACLOUD (S1) and SOCRATES (S2) but then combined for both campaigns in the histogram plot of the percentages of rimed particles per ice particle habit. – Please explain?

To avoid confusion, the dependent frequency of occurrence distribution of different ice particle habits was combined for SOCRATES and ACLOUD (see answer to major comment 8).