**Reply to reviewer #1**

**General comments:**

This manuscript is a novel attempt to assimilate C-band polarimetric microphysical retrievals into a convection-resolving model in Germany compared to the legacy assimilation approach using Z. The authors find generally positive benefits, particularly for LWC, with IWC neutral/beneficial in stratiform cases but with harmful effects in a convective case. This is not particularly surprising, as the retrievals used are not formulated for rimed ice/hail, but it is nonetheless a good demonstration of what can happen if the retrieval equations are not applied appropriately/selectively to how they were formulated. The results overall indicate that the addition of LWC and IWC retrievals, with Z, results in the best forecasts, and encourages further exploration. The manuscript is very well written, with a thorough introduction section that clearly establishes the state-of-the-art of polarimetric radar DA and its challenges. The basis and results of the study are novel and timely as polarimetric radars are increasingly implemented and their information content explored. There are a number of improvements I'd like to see to the manuscript primarily regarding clarity or requests for additional context/information. While I don't think it is necessary for publication, I did feel the manuscript could benefit from one or two plots actually depicting the results of the different assimilation experiments (e.g., actual post-assimilation hydrometeor and/or moisture fields, or QPF fields, vs. the observed QPE; the distribution/statistics of the IWC/LWC fields, etc.), so that readers can get a visual sense of the effects these are having rather than solely relying on only statistics. Despite this, I believe the manuscript will be ready for publication once the following minor comments are addressed.

*Dear reviewer, thank you very much for the positive feedback and encouraging words. We appreciate your efforts and your suggestion to include visual impressions of the effect of the LWC/IWC assimilation. We indeed already tested the visualization of e.g. QPF vs. QPE fields when LWC/IWC assimilation was included compared to the case when only Z was assimilated. However, while some slight positive impact could be noted in terms of FSS, BSS and FBI, it was very difficult to identify the differences visually. Figure A shows an example for a single observation experiment we made. While some effect of single Z or LWC/IWC observations on the QPF fields can be noted, the differences in the fields between the Z and LWC/IWC assimilation is difficult to identify visually. Therefore, we decided to not visualize the QPF/QPE fields in the paper and to just focus on the statistics of the FSS, BSS and FBI measures. Still, a deeper look into the impact of the LWC/IWC assimilation on the spatial statistics of the LWC, IWC or moisture fields could be worth looking at in the future. Thanks again for your helpful comments and suggestions!*

*Note that all new line numbers in the following refer to the manuscript file without tracked changes!*

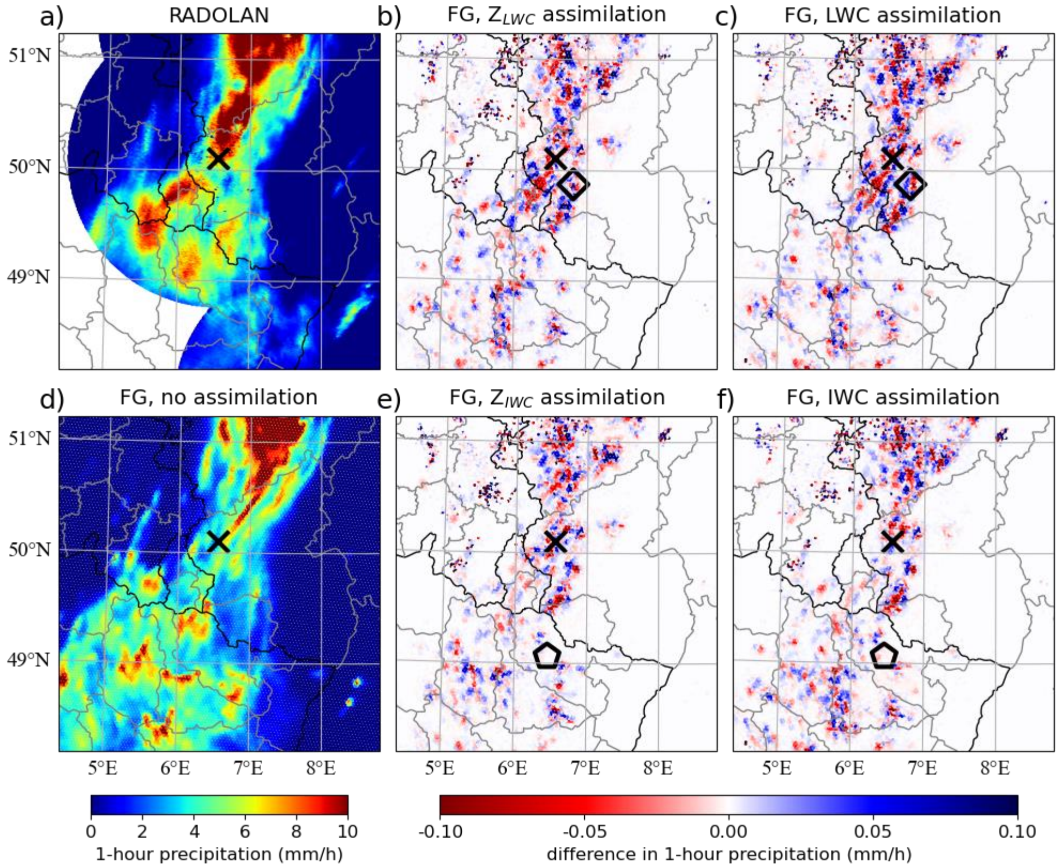


*Figure A: a) 1-hour RADOLAN precipitation accumulation for 14 July 2021 16 UTC, d) first guess (FG) 1-hour precipitation for the same hour without data assimilation, b) difference to d) for assimilation of a single Z super observation at the position below the melting layer marked by the black diamond, c) same as c) but with a single LWC super observation, e) and f) are as b) and c) but for a superobbing point above the melting layer marked by the black pentagons.*

**Specific comments:**

1. L26: Please define FSS and FBI for the abstract.

   *Of course, thank you! (L26,27)*

2. L103: It may be worth mentioning that beyond the rudimentary treatment of PSDs, hydrometeor shapes and orientations are rarely (if ever?) taken into account at all within model microphysics schemes.

   *Thank you for that comment. We adjusted the text accordingly. (L104)*

3. L160: Which microphysics scheme is used in the ICON-D2? A citation is needed here.

   *We included the corresponding citation (Doms et al., 2011). (L161)*

4. L180: In this study, when Z is assimilated (either in CNV+Z, CNV+[LWC + IWC]/Z, or CNV + LWC + IWC + Z) does this mean it is both directly assimilated *and* assimilated through LHN? It is not explicitly clear to me how Z is being used as the reference assimilation study. (Edit: I see now on L338 it is stated that LHN is excluded in this case, but this wasn't clear at first when mentioned in the context of the operational Z assimilation scheme).

*We included the sentence “Note that LHN and the assimilation of 3D V observations are not applied in this study (see below).” for clarity. (L183,184)*

5. L204 and elsewhere: Is there a specific reason the variables are included as ZDR, PHIDP, KDP, and RHOHV rather than $Z_{DR}$, $\Phi_{DP}$, $K_{dp}$, and $\rho_{hv}$? It isn’t critical, but I think use of the variables are more standard notation, would clean up the equations/discussion, and not suggest that ZDR, RHOHV, PHIDP, and KDP might be acronyms to unfamiliar readers.

   *This is a good point. There is no real reason why we wrote the variables like that. We changed the text to the standard notation everywhere. (e.g. L206-208)*

6. L206: Strictly speaking, isotropic scatterers will have an intrinsic ZDR of exactly 0 dB rather than one close to 0 dB.

   *Of course! (L209)*

7. L217: This value of $\rho_{hv}$ seems much lower than what I would have expected. At S band, $\rho_{hv}$ in dry snow/ice (rather than melting particles, which are neglected in these retrievals) is usually nowhere near 0.85 unless it is very large hail experiencing resonance scattering, in which case the use of the IWC retrieval equations would be inappropriate anyway (both due to it being hail rather than snow and due to the non-Rayleigh scattering). With all the tests done to find the optimal DAP, were any tests done to examine the impact of these thresholds and find optimal values? What did the typical distribution of $\rho_{hv}$ values actually look like above the ML? I am concerned that using such a low $\rho_{hv}$ threshold aloft will necessarily retain data that is going to have very noisy $\Phi_{DP}$/$K_{dp}$ and thus noisy/erroneous IWC retrievals that should not be assimilated. I would be curious if a much more stringent $\rho_h$ threshold (perhaps something like 0.92 or greater) would result in better retrieval/assimilation results.

   *This is a legal comment. We chose the RHOHV threshold just based on the value listed in the Kumjian (2013a) paper for S-band. Yes, possibly this too low threshold contributes to the rather worse assimilation results for IWC by retaining noisy data. However, we did not test the sensitivity of the assimilation experiments to the RHOHV thresholds. This would be worth an investigation in the future. Besides, you find a typical distribution of RHOHV and Z in our data in Figures B and C.*

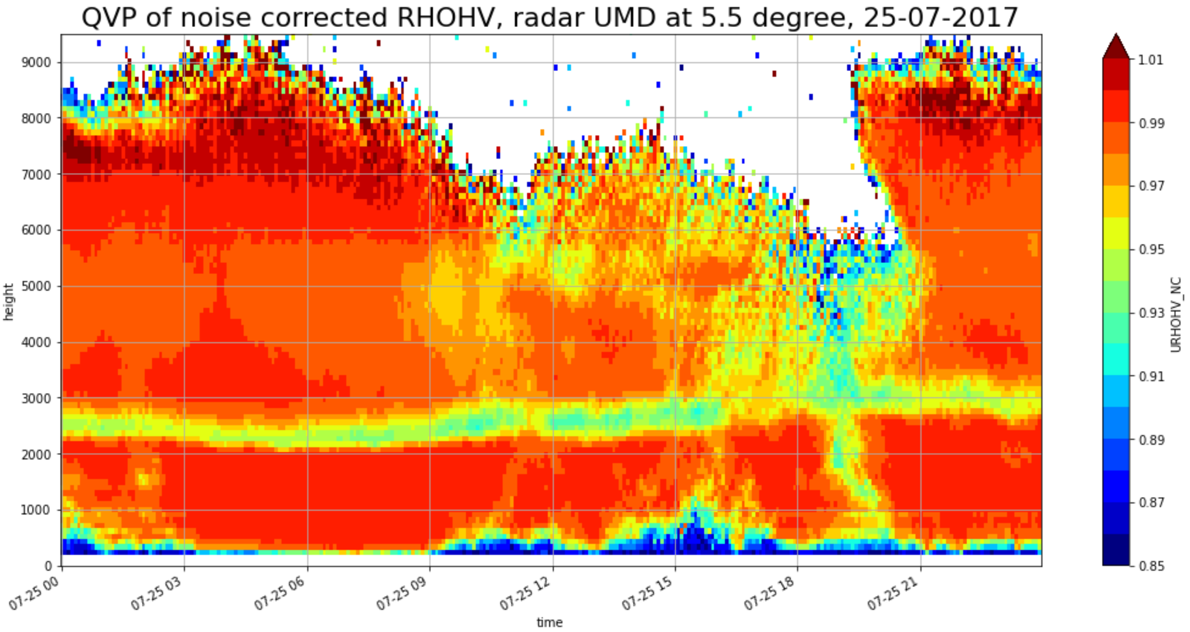


*Figure B: QVP of RHOHV of radar UMD at 5.5 degrees elevation for the stratiform day 25 July 2017.*

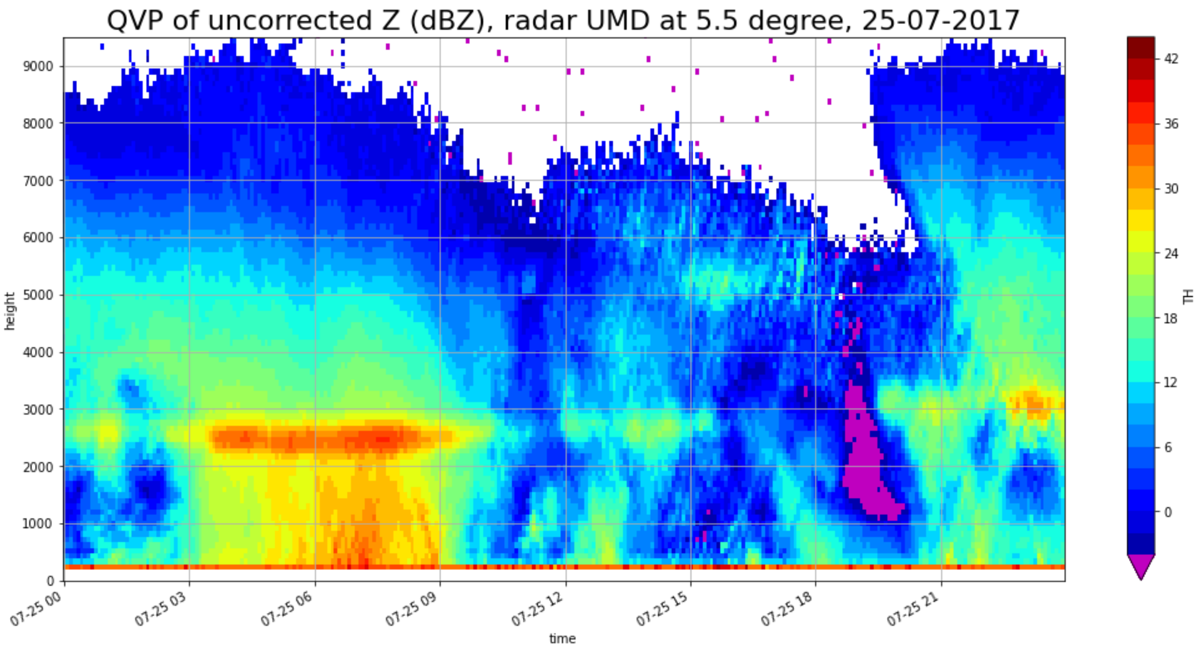


*Figure C: Corresponding QVP of Z (in dBZ).*

8. L221: How is the bottom of the melting layer determined, as shown in Fig. 2? I assume some sort of $\rho_{hv}$ threshold when using a QVP, but for the convective case (for which I presume model data was relied upon) it is less clear.

   *We used QVPs of RHOHV, determined the approximate center of the melting layer within 6-hour intervals visually and subtracted/added 500m to identify the approximate upper and lower boundaries. In convective situations, we used the closest operational radio sounding to approximate the height of the center (approx.. 0°C isotherm) of the melting layer and then added/subtracted 500m to obtain the upper/lower boundaries. We changed the text to: "The height of the melting layer is determined from so-called Quasi-Vertical Profiles (i.e., azimuthal medians of PPIs measured at sufficiently high* *elevations and transferred to range-height displays; Trömel et al., 2014; Ryzhkov et al., 2016), as derived from PPIs measured at a 5.5 degree elevation angle, or from the nearest operational DWD radio sounding, especially in convective situations." (L227)*

9. L225: I appreciate why such a large window is needed for consistency, but depending on how small and isolated the convection is, could 9 km be too large a window and end up heavily incorporating data boundary regions into the Kdp calculation for precipitation regions?

   *Yes, this is a big problem with the low radial resolution of most of the radar data used. On the one hand side, a small window size close to strong cells is desireable to keep the small scale features, but on the other hand side, the low resolution may introduce errors when using a small window (because only little data points contribute to the estimates). The newer radar data of DWD (since march 2021) has a 4 times higher radial resolution (250m) compared to the older data (1 km). Thus, it would be worth exploring how an adjustment of the window size for KDP depending on the precipitation situation may change the assimilation results. Unfortunately, the convective case considered is from 2017 and thus has the low 1-km resolution. For sure we will keep on working on the topic including the relevance of the enhanced radial resolution and hopefully the assimilation of polarimetric microphysical retrievals will become operational at DWD in the future. However, this is beyond the scope of this study.*

10. L246 and elsewhere: I assume log here is $\log_{10}$ and not ln? It may be helpful just to clarify.

    *Yes, this is a bit unclear. We changed "log" to "log10" everywhere. (e.g. L249)*

11. L256: It may also be worth mentioning in addition to just hail that R(A) at C band struggles from the resonance scattering of medium-large-sized raindrops, which causes R(Kdp) to outperform R(A) for moderate to heavy rainrates.

    *Thanks for the comment. We added one sentence: "In addition, resonance scattering of medium and large sized raindrops at C-band may favour the use of LWC(KDP) compared to LWC(A) in moderate to heavy rain." (L259-261)*

12. L207: I am a bit confused by this. The Figure 2 caption says LS = 10 km, which would make it equivalent to LC, but here it makes it sound like LS != LC.

    *We hope it becomes more clear by adding the word "also": "The number of radar bins contributing to the averaging decreases with increasing distance from the radar, and the window size for the averaging (winsize_avg in km) is equal to res_cartesian in KENDA, but is also modified in our study while keeping res_cartesian constant." (L293) Note that the acronyms "LC" and "LS" are changed here to "res_cartesian" and "winsize_avg", respectively, for reasons of clarity.*

13. L305: While the results of these tests are not shown here, previous studies such as Liu et al. (2020) have demonstrated the same thing for hydrometeor mixing ratios and could be cited here, if desired.

Liu, C., M. Xue, and R. Kong, 2020: Direct variational assimilation of radar reflectivity and radial velocity data: Issues with nonlinear reflectivity operator and solutions. Mon. Wea. Rev. **148**.

*Thanks for the suggestion. We added the citation! (L311)*

14. L331: I may be misunderstanding something simple, but if the assimilation process during this period is what is done operationally (CONV + Z + LHN), what is the reason the data was obtained 24-35 hours beforehand and further spun up rather than just using the operational data from the model initial time (e.g, 00 UTC 13 July 2021 instead of 00 UTC 12 July 2021)?

    *Good point. The problem is a very limited access to the DWD data. The model data was only provided for the times 00 UTC 12 July 2013 etc. while the time interval of interest only started 24 etc. hours later. To reduce confusions, we rephrased that the model data was provided for the initial times of the experiment periods, which is reasonable because the assimilation cycles used to obtain the model data at the starting times of the experiments are identical to the operational cycles. We rewrote the paragraph 4.3 accordingly. (L334-338)*

15. L336: I am curious about the assimilation of Z data only within the melting layer, where the relationship between Z and the microphysical state variables becomes most complicated and obfuscated. If anything the data in the melting layer is often neglected because it can result in some large errors.

    *That is true. However, in the operational ICON-D2 routine at DWD, Z is assimilated within the melting layer. Since the operational Z assimilation (without LWC or IWC) is used as reference in this paper, the configurations assimilating LWC/IWC also need to include the Z assimilation in the melting layer for comparability. However, as indicated by the reviewer, results of assimilating Z data in the melting layer highly depend on the operator's melting scheme. Different flavours of the Maxwell-Garnett-, Bruggemann- and Wiener Effective Medium Approximations (EMA) can be chosen to explore the uncertainty in the melting layer. We are currently investigating deficiencies of the simulated melting layer signature, parts of which may be attributed to the fact that in the model microphysics parameterization, meltwater from snow, graupel, and ice is instantaneously shedded into the rain class, causing too small and too few remaining frozen particles in the melting layer. The current version of EMVORADO estimates a melted fraction as function of temperature and particle size (Blahak 2016) as part of the remaining frozen mass without "back-shuffling" some rain water to the particles. This leads to a systematic underestimation of the melting effect in all radar moments, despite the quite detailed consideration of various EMAs for the effective refractive index. We will explore better approaches, e.g. a wet snow class borrowing parts of the rain and snow mixing ratios and mix them (e.g. Jung et al. 2008a, 2008b, 2010), to reduce the bias to the observations. However, as soon as model microphysics with explicit mixed-phase snow, graupel, and hail (e.g, as in Frick et al. 2013 for snowflakes) becomes available in the future, their liquid fraction could be directly applied in the forward operator.*


    *Blahak, U.: RADAR_MIE_LM and RADAR_MIELIB – Calculation of Radar Reflectivity from Model Output, COSMO Technical Report 28, Consortium for Small Scale Modeling (COSMO), available at: http://www.cosmo-model.org/content/model/documentation/techReports/cosmo/docs/techReport28.pdf (last access: 10 January 2022), 2016. a, b, c, d, e*

    *Jung, Y., G. Zhang, and M. Xue, 2008a: Assimilation of simulated polarimetric radar data for a convective storm using the ensemble Kalman filter. Part I: observation operators for reflectivity and polarimetric variables. Monthly Weather Review, 136 (6), 2228-2245, DOI: 10.1175/2007MWR2083.1.*

    *Jung, Y., M. Xue, G. Zhang, and J. M. Straka, 2008b: Assimilation of simulated polarimetric radar data for a convective storm using the ensemble Kalman filter. Part II:*

*impact of polarimetric data on storm analysis. Monthly Weather Review, 136 (6), 2246-2260, DOI: 10.1175/2007MWR2288.1.*

*Jung, Y., M. Xue, and G. Zhang, 2010: Simulations of polarimetric radar signatures of a supercell storm using a two-moment bulk microphysics scheme. Journal of Applied Meteorology and Climatology, 49 (1), 146-163, DOI: 10.1175/2009JAMC2178.1.*

*Frick, C., Seifert, A., and Wernli, H.: A bulk parametrization of melting snowflakes with explicit liquid water fraction for the COSMO model, Geosci. Model Dev., 6, 1925–1939, https://doi.org/10.5194/gmd-6-1925-2013, 2013.*

16. L512: I apologize if I'm misunderstanding, but these two sentences seem contradictory to me. It is stated that the best forecasts are achieved when there is limited influence from the radar-based retrievals (I assume in terms of impact to the analysis by way of a larger lower threshold, rather than spatial localization?). But subsequently it is stated that a smaller observation error standard deviation, which I believe would enhance the weighting toward observations, is also beneficial.

    *Indeed, this may be confusing. We changed the first sentence as follows: "Thus, best first guess of precipitation forecasts are achieved when the influence of the observed microphysical estimates on the model state is rather small in terms of observation localization length-scale and lower data limit. A rather small observation error standard deviation of 0.25 in log10(LWC) and log10(IWC) was most successful." (L523-527)*

17. L548: This, to me, is an absolutely crucial piece of information for interpreting the results of assimilating LWC/IWC in this study and needs to be discussed earlier in the paper in the assimilation section. While I think the adjustment of the moisture variables (rather than precipitation variables) is in general an important aspect of radar data assimilation, I also feel the impacts of assimilating a "bad" retrieval (say above the ML in what is actually graupel/hail) on moisture may be even more deleterious than if it were acting on qs, which could precipitate out relatively quickly and lead to little harm to the forecast. This seems like an important follow-up study since it is likely not what readers would have expected at first.

    *This is a very interesting point. Following this idea, producing increments also in qs when assimilating IWC would potentially not change the results much. This definitely requires future investigation. We included one sentence in section 4.2 in order to enable a better interpretation of the results: "Moreover, microphysical analysis increments of only cloud water mixing ratio and specific humidity are produced, i.e., not all available hydrometeor species (e.g., rain, cloud ice, and graupel mixing ratios) are updated individually in KENDA's standard configuration." (L.320-322)*

18. L565: Could additional studies be done that only assimilate the IWC retrievals in areas identified as snow in convective cases, thus potentially limiting the presumed harmful impacts of assimilating poor/inappropriate retrievals of IWC? With benefits coming from Z alone, I am curious if it could still be wise to assimilate IWC only selectively, in addition to LWC and Z (and V and CNV).

    *Thank you for this input! A hydrometeor classification prior to assimilation may have some potential to eliminate the harmful effects of hail/graupel in convective precipitation. Also, a selective assimilation of IWC in such situations could be worth being tested.*

**Technical corrections**:

1. L31: “adequat” should be “adequate”
2. L330: “and including LHN” should be “including LHN”
3. L538: “adjusted” should be “developed”

   *Points 1) and 3) were corrected (L31 and L541), point 2) was not corrected because the respective sentence was deleted. Thank you again for your time!*