Evaluation of polarimetric ice microphysical retrievals with OLYMPEX campaign data

Armin Blanke¹, Andrew J. Heymsfield², Manuel Moser³,⁴, and Silke Trömel¹,⁵

¹Department of Meteorology, Institute of Geosciences, University of Bonn, Bonn, 53121, Germany
²Mesoscale and Microscale Meteorology Lab (MMM), National Center for Atmospheric Research, Boulder, Colorado, 80307, USA
³Institute for Physics of the Atmosphere, University Mainz, Mainz, 55099, Germany
⁴Institut für Physik der Atmosphäre, Deutsches Zentrum für Luft- und Raumfahrt, Oberpfaffenhofen, 82234, Germany
⁵Laboratory for Clouds and Precipitation Exploration, Geoverbund ABC/J, Bonn, 53121, Germany

Correspondence: Armin Blanke (armin.blanke@uni-bonn.de)

Received: 22 December 2022 – Discussion started: 13 January 2023
Revised: 14 March 2023 – Accepted: 15 March 2023 – Published: 20 April 2023

Abstract. Polarimetric microphysical retrievals reveal a great potential for the evaluation of numerical models and data assimilation. However, the accuracy of ice microphysical retrievals is still poorly explored. To evaluate these retrievals and assess their accuracy, polarimetric radar measurements are spatially and temporally collocated with in situ aircraft measurements obtained during the OLYMPEX campaign (Olympic Mountain Experiment). Retrievals for ice water content (IWC), total number concentration $N_t$, and mean volume diameter $D_m$ of ice particles are assessed by comparing an in situ dataset obtained by the University of North Dakota (UND) Citation II aircraft with X-band Doppler on Wheels (DOW) measurements. Sector-averaged range height indicator (RHI) scans are used to derive vertical profiles of microphysical retrievals. The comparison of these estimates with in situ data provides insights into strengths, weaknesses, and the accuracy of the different retrievals and quantifies the improvements in polarimetry-informed retrievals compared to conventional, non-polarimetric ones. In particular, the recently introduced hybrid ice water content retrieval exploiting reflectivity $Z_H$, differential reflectivity $Z_{DR}$, and specific differential phase $K_{DP}$ outperforms other retrievals based on either ($Z_H, Z_{DR}$) or ($Z_H, K_{DP}$) or non-polarimetric retrievals in terms of correlations with in situ measurements and the root mean square error.

1 Introduction

Polarimetric microphysical retrievals bear great potential for data assimilation and the evaluation of numerical models, however, their exploitation is still in its infancy. For instance, Trömel et al. (2021) demonstrated the potential of using polarimetric observations and retrievals to evaluate and improve microphysical parameterizations. Pioneering work by Carlin et al. (2016) revealed the benefits of assimilating polarimetric microphysical retrievals. Similar work is currently underway in Germany. Reimann et al. (2021) took a first step towards assimilating polarimetric variables into the ICOsahedral Nonhydrostatic (ICON) model (Zängl et al., 2015) via adapting microphysical retrievals for their application to observations of the polarimetric C-band radar network of the German national meteorological service.

Ryzhkov et al. (1998) pointed to the limited database to identify the main reasons for the differences between individual in situ measurements and polarimetric retrievals. In fact, the in-depth evaluation of retrievals requires extensive airborne in situ cloud particle measurements over polarimetric radar sites using, e.g., so-called optical array probes (OAPs) collected during field campaigns. However, these are substantial high budget and provide data only along flight trajectories. Hogan et al. (2006) introduced an ice water content (IWC) based on radar reflectivity $Z$ and atmospheric temperature $T$, Tian et al. (2016) evaluated its performance along with that of a mean volume diameter $D_m$ retrieval with air-
craft in situ data from the Bow Echo and Mesoscale Convective Vortex Experiment (BAMEX; Davis et al., 2004). They observed an overestimation of the mean Hogan IWC retrieval compared to in situ measurements (1.52 g m⁻³ vs. 1.25 g m⁻³) and a correlation of 0.55. Similarly the mean retrieved $D_m$ showed an overestimation (2.08 mm vs. 1.77 mm) and a low correlation of 0.27.

One reason for the bad performance of non-polarimetric retrievals is that the horizontal reflectivity $Z_H$ in snow is approximately proportional to the fourth moment of the particle size distribution (PSD) (Hu and Ryzhkov, 2022); hence $Z_H$ is insensitive to small particles, whereas other moments, such as the IWC, are sensitive to the small particle contributions. In contrast, specific differential phase $K_{DP}$ is proportional to the first moment of the PSD and thus the whole spectrum is considered. However, $K_{DP}$ strongly depends on the aspect ratio and orientation of the particles, necessitating prior knowledge of these parameters. Aydin and Tang (1997) proposed for IWC estimation the combination of $K_{DP}$ and differential reflectivity $Z_{DR}$ because their ratio is not affected by the variability in orientation and particle aspect ratio (Ryzhkov et al., 2018). Another set of polarimetric relations to quantify snow properties was derived by Bukovčić et al. (2018) exploiting $Z_H$ and $K_{DP}$. More recently, Carlin et al. (2021) suggested a hybrid application for estimating IWC by combining the complementary strengths and optimal ranges of the IWC retrievals following Bukovčić et al. (2018) and Ryzhkov and Zrnić (2019). Apart from IWC retrievals, polarimetric retrieval relations have been suggested for $D_m$ and total number concentration of ice particles per unit volume $N_t$ based on approaches utilizing combinations of three or two polarimetric variables, always including $K_{DP}$ (Ryzhkov et al., 2018; Bukovčić et al., 2020). To evaluate the quality of polarimetric retrievals, approaches based on in situ and/or ground-based measurements were pursued.

Nguyen et al. (2019) proposed a methodology to retrieve IWC using $Z_{DR}$ and $K_{DP}$ from X-band dual-polarization airborne radar data. This algorithm was found to be superior to power-law fits using $Z$ compared to others because $Z_{DR}$ minimizes the dependence of IWC on variations in ice particle shape and orientation. An evaluation with in situ data from the High Altitude Ice Crystal – High Ice Water Content (HAIC-HIWC) field campaign revealed that the additional use of $Z_{DR}$ reduced the root mean square difference by 6% and the bias by 15% on average compared to retrievals using $K_{DP}$ only.

Several noise-reducing techniques for reconstructing average vertical profiles from radar data have been proposed and used in the literature (Table 1). They are based on plan position indicator (PPI) or range height indicator (RHI) scans using single or multiple elevations. However, polarimetric and especially the phase-based radar measurements may be noisy in ice and snow and even more near the cloud top. As a consequence, it is beneficial to reduce their statistical noise before the calculation of microphysical retrievals. For instance...
adjustments of

Ryzhkov and Zrnić (2019) and verified whether the expected

tative Archive Center (DAAC). As a cross-check, we followed

Global Hydrology Resource Center (GHRC) Distributed Ac-

(OLYMPEX/DOW/DATA201), were downloaded from the

study, we exclusively utilize the latter frequency as only

measurements of the lower frequency were available after

of 12 November 2015. More detailed information on the DOW

can be found in Houze et al. (2018).

The DOW radar (Wurman et al., 1997), placed within

Chehalis Valley at Lake Quinault (47.48° N, 123.86° W; at

64 m altitude), Washington, and operated by the Center of

Severe Weather Research (CSWR), is installed on a mobile

truck. The polarimetric dual-frequency X-band radar oper-

ated with a range of 59.96 km and a radial resolution of

75 m and used two independent transmitters at frequencies

of ∼ 9.55 GHz and ∼ 9.40 GHz (Houze et al., 2018). In this

study, we exclusively utilize the latter frequency as only

measurements of the lower frequency were available after

12 November 2015. More detailed information on the DOW

can be found in Houze et al. (2018).

The DOW’s 10 min scanning schedule includes plan posi-
tion indicator (PPI) scans for the azimuthal sector be-

between 39.2 and 83.6° at six elevations between 2.8 and

11° and a series of 22 range–height indicators (RHIs) for

the azimuthal sector between 50.4 and 71.4° in equidistant

intervals of 1° and elevations ranging from 0 to

71°. For the evaluation of polarimetric ice microphysi-
cal retrievals, we focus in this study on the sector RHIs.

All datasets utilized (Petersen et al., 2018), including a

new version of the DOW data with improved calibration

(Houze et al., 2018; doi: https://doi.org/10.5067/GPMGV/

OLYMPEX/DOW/DATA201), were downloaded from the

Global Hydrology Resource Center (GHRC) Distributed Ac-
tive Archive Center (DAAC). As a cross-check, we followed

Ryzhkov and Zrnić (2019) and verified whether the expected

ZH–ZDR relationship for X-band in rain is fulfilled and made

adjustments of ZDR if necessary. Accordingly, a ZDR correc-
tion was applied to all data used, with a larger correction re-
quired for measurements in November than in December (on

average 0.38 dB vs. 0.19 dB).

Specific differential phase KDp, defined as half the range

derivative of differential propagation phase shift ΦDP, is es-
timated following Vulpiani et al. (2012). For processing

efficiency, the derivative of ΦDP is approximated using low-

noise Lanczos differentiators (Holoborodko, 2008). Since

KDp is inversely proportional to the radar wavelength, more

reliable estimates can be expected at X-band compared to

C- or S-band with according benefits also for microphysical

retrievals. Only data with a cross-correlation coefficient ρhv

above 0.7 are used for the KDp estimation in order to reduce

the impact of noisy and non-meteorological contamination.

2.1 Polarimetric data from X-band radar DOW

The in situ microphysical cloud measurements during the

OLYMPEX campaign have been acquired with the Univer-
sity of North Dakota (UND) Cessna Citation II airplane,

which was equipped with an enhanced instrumental pay-

load, including the 2D stereo imaging probe (2D-S; SPEC

Inc, USA) and the high-volume precipitation spectrometer

(HVPS; SPEC Inc, USA). Both instruments provide shad-

owgraphs of cloud particles. The 2D shadow images of hy-
drometeors are generated as the particles penetrate through

the sampling area of the particle imagers. A detailed de-
scription of the 2D-S and HVPS operating principles, un-
certainties, and limitations can be found in Lawson et al.

(2006), Baumgardner et al. (2017), and Moser et al. (2023).

Note that the data recorded by the 2D-S and HVPS differ

in sampling volume size and pixel resolution. The 2D-S is

equipped with 128 pixels of 10 µm resolution each, which

allows imaging particles from 10 µm to 1.28 mm in size. With

the larger pixel resolution provided by the HVPS of 150 µm,

imaging of hydrometeors from 150 µm to 3.25 cm is enabled.

Here, the datasets of the horizontally oriented HVPS and the

horizontally oriented part of the 2D-S processed by Heyms-

field et al. (2018) were combined with a switch-over size at

1000 µm; i.e., particles smaller than 1000 µm are sized by the

2D-S and larger particles by the HVPS. Standard process-
ing and correction options with SODA (software for OAP

data analysis, provided by Aaron Bansemer, National Cen-
ter for Atmospheric Research/University Corporation for At-
mospheric Research; Bansemer, 2023) were applied includ-
ing shattering and dead-time corrections. For a better com-
parability of the in situ measurements with the according

radar-based retrievals of IWC, Dw, and Nr, only particles

larger than 100 µm are considered because of the sensitiv-

Table 1. Overview of areal averaging techniques to derive quasi-vertical profiles of polarimetric radar variables.

<table>
<thead>
<tr>
<th>Methodology</th>
<th>Acronym</th>
<th>Scan strategy</th>
<th>Used azimuth</th>
<th>Used elevations</th>
<th>Citation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Quasi-vertical profile</td>
<td>QVP</td>
<td>PPI</td>
<td>360°</td>
<td>Single high</td>
<td>Ryzhkov et al. (2016)</td>
</tr>
<tr>
<td>Range-defined QVP</td>
<td>RD-QVP</td>
<td>PPI</td>
<td>360°</td>
<td>Multiple</td>
<td>Tobin and Kumjian (2017)</td>
</tr>
<tr>
<td>Columnar vertical profile</td>
<td>CVP</td>
<td>PPI</td>
<td>Sector</td>
<td>Multiple</td>
<td>Murphy et al. (2020)</td>
</tr>
<tr>
<td>Enhanced vertical profile</td>
<td>EVP</td>
<td>PPI</td>
<td>Sector</td>
<td>Multiple</td>
<td>Bukovič et al. (2017)</td>
</tr>
<tr>
<td>Slanted vertical profile</td>
<td>SV</td>
<td>PPI</td>
<td>Sector</td>
<td>Single low</td>
<td>Bukovič et al. (2017)</td>
</tr>
<tr>
<td>Range height indicator QVP</td>
<td>R-QVP</td>
<td>RHI</td>
<td>Fixed</td>
<td>Multiple</td>
<td>Allabakash et al. (2019)</td>
</tr>
</tbody>
</table>

ity of weather surveillance radars (Ryzhkov et al., 2020) and uncertainties in probe data (Poellot and Bansemer, 2017; Baumgardner et al., 2017). With the particle size distribution (PSD) given by Heymsfield et al. (2018), $D_m$, IWC, and $N_i$ for particles between 100 $\mu$m and 3 cm are calculated. In order to derive $N_i$, the number concentration for each particle size bin is added up.

The different relations between measures of particle sizes used in the radar community and those derived from in situ measurements must be considered in their comparison. Radar-based retrievals mostly provide the mean volume diameter $D_m$ as a parameter for the size of the particles, which is defined as the ratio of the fourth to the third moment of the PSD:

$$D_m = \frac{\int D^4 N(D)dD}{\int D^3 N(D)dD},$$  

(1)

where $D$ is the equivolume diameter. It should be noted that $D_m$ is very close to the median volume diameter $D_0$. Assuming that the PSD follows a gamma distribution with shape parameter $\mu$ (Ulbrich, 1983), we obtain

$$D_m = \frac{4 + \mu}{3.67 + \mu} D_0.$$  

(2)

Instead of $D_m$, the aircraft microphysical probes usually measure either the median mass diameter $D_{mm}$ or the median volume diameter $D_{mv}$ of the distribution of maximal particle dimension (Hu and Ryzhkov, 2022), which is given by

$$D_{max} = D\phi^{-1/3},$$  

(3)

where $\phi$ is the particle aspect ratio. For a gamma size distribution and taking Eq. (3) into account, we yield

$$D_{mm} = \frac{2.67 + \mu}{4 + \mu} D_m \phi^{1/3},$$  

(4a)

$$D_{mv} = \frac{3.67 + \mu}{4 + \mu} D_m \phi^{1/3},$$  

(4b)

with Eq. (4a) assuming particle density being inversely proportional to its size. Assuming $\phi = 0.6$ and an exponential distribution ($\mu = 0$), the following two $D_m$ relationships (Hu and Ryzhkov, 2022) are obtained:

$$D_m \approx \frac{D_{mm}}{0.79}$$  

(5a)

and

$$D_m \approx \frac{D_{mv}}{1.09}.$$  

(5b)

These relations enable the direct comparison between the in situ measurements with the radar-based $D_m$ retrievals, with the median mass size $D_{mm}$ derived in situ utilized in our study to obtain $D_m$. The IWC is estimated by using a mass-dimension relation between mass $m$ and size given by

$$m = aD_{max}^b,$$  

(6)

with $a = 0.0121$ kg per m$^b$ $b = 1.9$, and $D_{max}$ the diameter of the minimum enclosing circle of the projected 2D image. We followed Chase et al. (2018) by adopting Eq. (6) and parameters $a$ and $b$ from Brown and Francis (1995) and modifying them considering the particle size definition by Hogan et al. (2012). Even though parameters $a$ and $b$ of the mass-dimension relationship vary with the environmental conditions and particle shapes (Baker and Lawson, 2006), constant standard parameters are used in this study which reasonably represent the mean ice water content, especially for ice crystal aggregates. Tridon et al. (2019) confirmed that the aggregates observed during OLYMPEX can mostly be described by a quite narrow range of mass-size relations. In single situations with large aggregates or intense riming processes, however, the fixed parametrization may underestimate the ice water content (see also Heymsfield et al., 2023).

In addition to the quantitative information provided by the 2D-S and HVPS, high-resolution in situ probe imagery data from the cloud particle imager (CPI; SPEC Inc, USA) can be used to accurately identify particle types and characteristics. Only with the CPI is it possible to directly monitor supercooled liquid water (SLW) droplets of micron size attached to ice particles and thus diagnose riming unambiguously and estimate the degree of riming.

The Rosemount icing detector (RICE; Baumgardner and Rodi, 1989) mounted on the Citation II is used as a supporting probe to detect the presence of SLW mandatory for riming (Vogel and Fabry, 2018). The RICE oscillates at a constant frequency, but when supercooled droplets freeze on its surface, the frequency of vibration decreases. Once accumulated ice exceeds a certain threshold, the probe tip is briefly heated to remove accreted ice. Data are available again as soon as the probe temperature has stabilized (Heymsfield and Miloshevich, 1989).

### 3 Radar-based microphysical retrievals

This section summarizes the most recent polarimetric and a suite of conventional non-polarimetric ice microphysical retrievals for IWC, $N_i$, and $D_m$ considered and assessed in this study. Two conventional $D_m$ retrievals derived from statistical relations between particle sizes and reflectivity expressed in linear scale ($Z_h = 10^{6.12}m$; in units of mm$^6$ m$^{-3}$) are used in our analysis. The first relation introduced by Skofronick-Jackson et al. (2019) is based on a power law between $D_m$ (mm) and Ku-band $Z_h$ fitted to data from the GPM Cold Season Precipitation Experiment (GCPEX) campaign (Skofronick-Jackson et al., 2015) conducted in Canada:

$$D_{m}^4 (Z_h) = 1.45 Z_h^{0.25}.$$  

(7)
Matrosov et al. (2019) introduced another power-law relation derived from $Z_h$ data of ground-based S-band radar and aircraft in situ calculated $D_{\text{mv}}$ obtained during the Indirect and Semi-Direct Aerosol Campaign (ISDAC) in Alaska (Maahn et al., 2015). Assuming $\psi = 0.6$ and using Eq. (5b), it follows that

$$D_{\text{mv}}^{H}(Z_h) = \frac{1}{1.09} \cdot \left(1.15Z_h^{0.271}\right)^{1/3}. \quad (8)$$

This equation differs from Eq. (5) in Murphy et al. (2020) showing, due to a typo, the reciprocal multiplier when converting $D_{\text{mm}}$ to $D_{\text{mv}}$.

Since $Z_h$-based $D_{\text{mv}}$ retrievals are disproportionally weighted by a few large particles, polarimetric $K_{\text{DP}}$-based retrievals have a great potential to provide more accurate estimates. Additionally, such estimators use the key advantage that $K_{\text{DP}}$ is not affected by attenuation and not biased by noise and radar miscalibration. To retrieve a polarimetric $D_{\text{mv}}$ from $K_{\text{DP}}$ and $Z_{\text{dp}}$, where $Z_{\text{dp}} = Z_h - Z_v$ is the reflectivity difference at horizontal and vertical polarization in linear scale, as proposed in Ryzhkov et al. (2018) and Ryzhkov and Zrnčič (2019), we use

$$D_{\text{mv}}(Z_{\text{dp}}, K_{\text{DP}}) = -0.1 + 2.0 \left(\frac{Z_{\text{dp}}}{K_{\text{DP}}^3}\right)^{1/2}, \quad (9)$$

where $K_{\text{DP}}$ is in degrees per kilometer, $Z_{\text{dp}}$ is in millimeters to the power of six per cubic meter (mm$^6$ m$^{-3}$), and the radar wavelength $\lambda$ is in millimeters. This estimator is largely immune to variations in ice particle orientation and shape but has the inherent deficiency of being impacted by the degree of riming; therefore it is supposed to be more appropriate for lower-temperature regions where riming is less likely. As an alternative, Bukovčič et al. (2018, 2020) use $Z_h$ and $K_{\text{DP}}$ to retrieve

$$D_{\text{mv}}(Z_h, K_{\text{DP}}) = 0.67 \left(\frac{Z_h}{K_{\text{DP}}^3}\right)^{1/3}. \quad (10)$$

Unlike $D_{\text{mv}}(Z_{\text{dp}}, K_{\text{DP}})$, however, this retrieval is not immune to the variability in particle orientation and shape, and a strength of this $D_{\text{mv}}(Z_h, K_{\text{DP}})$ estimate is that it does not depend on density and is therefore not affected by the degree of riming.

Similarly to the aforementioned $D_{\text{mv}}$ retrievals, IWC retrievals can also be derived in a purely empirical fashion through the utilization of power laws. Hogan et al. (2006) introduced an expression which related in situ measured IWC (in g m$^{-3}$) to reflectivity (in dBZ) at various frequencies (e.g., 3 GHz) and temperature $T$ (in $^\circ$C) in the European Cloud Radiation Experiment (EUCREX) derived in the Rayleigh approximation:

$$\log_{10}(\text{IWC}(Z_h, T)) = 0.06Z_h - 0.0197T - 1.7. \quad (11)$$

They also exploited the IWC$^1(Z_h, T)$ to evaluate the mesoscale version of the Met Office Unified Model. However, the IWC relationship implicit in the models’ parameterization (see Hogan et al., 2006, for details) is

$$\log_{10}\left(\text{IWC}^{II}(Z_h, T)\right) = 0.06Z_h - 0.0212T - 1.92. \quad (12)$$

Nguyen et al. (2019) proposed two more empirical but polarimetric IWC retrievals via optimal fitting parameters, exploiting aircraft measurements from a polarimetric side-pointing X-band radar and measured IWC by in situ probes obtained during HAIC-HIWC. The two retrievals either utilize $K_{\text{DP}}$ only or include additionally the differential reflectivity $Z_{\text{dr}}$ expressed in linear scale ($Z_{\text{dr}} = 10^{0.1Z_{\text{dr}}}$):

$$\text{IWC}(K_{\text{DP}}) = 0.903K_{\text{DP}} + 0.319 \quad (13)$$

$$\text{IWC}^{1}(Z_{\text{dr}}, K_{\text{DP}}) = \frac{0.136K_{\text{DP}} + 0.037}{1 - Z_{\text{dr}}^{-1}}. \quad (14)$$

$Z_{\text{dr}}$ is set to 1.15 (Ryzhkov et al., 1998) when $Z_{\text{dr}}$ falls below this threshold. Ryzhkov et al. (1998) also demonstrated that $K_{\text{DP}}$ is sensitive to the aspect ratio $\psi$ and orientation of the particles, whereas IWC is not, requiring additional knowledge about the particles. Accordingly, the estimator $\text{IWC}(K_{\text{DP}})$ is highly affected by variations in $\psi$ and/or orientation, whereas the inclusion of $Z_{\text{dr}}$ in $\text{IWC}^1(Z_{\text{dr}}, K_{\text{DP}})$ reduces these dependences. The observational study by Nguyen et al. (2019) demonstrated that their empirical relation $\text{IWC}^1(Z_{\text{dr}}, K_{\text{DP}})$ is very close to the theoretical IWC relation by Ryzhkov et al. (2018):

$$\text{IWC}^{II}(Z_{\text{dr}}, K_{\text{DP}}) = 4 \times 10^{-3} \left(\frac{K_{\text{DP}}}{1 - Z_{\text{dr}}^{-1}}\right)^{1/3}. \quad (15)$$

The latter exploits the inherent information of $Z_{\text{dr}}$ about the shape and the orientation of particles. Similar to Eq. (9), $\text{IWC}^{II}(Z_{\text{dr}}, K_{\text{DP}})$ is practically insensitive to the shape and orientations of the ice particles because the numerator and denominator are proportionally impacted and thus the ratio is not affected.

In the absence of a birdbath scan in the scan schedule (as is, e.g., the case for the operational radar networks in the USA), ensuring high $Z_{\text{dr}}$ accuracy is often difficult. Bukovčič et al. (2018, 2020) introduced a generalized relation for IWC,

$$\text{IWC}(Z_h, K_{\text{DP}}) = \frac{10.2 \times 10^{-3}}{(F_0F_S)^{0.66}} (K_{\text{DP}})^{0.66} Z_h^{0.28}, \quad (16)$$

which is immune to $Z_{\text{dr}}$ miscalibrations and where $F_0$ is the orientation factor as a function of the width of the canting angle distribution $\sigma$ and $F_S$ the shape factor determined by $\psi$. The relation in Eq. (16) produces

$$\text{IWC}(Z_h, K_{\text{DP}}) \approx 0.31 K_{\text{DP}}^{0.66} Z_h^{0.28}, \quad (17)$$

for $\sigma = 0^\circ$, $\psi = 0.65$, and $\lambda = 32$ mm. In contrast to Eq. (10), however, the IWC$(Z_h, K_{\text{DP}})$ retrieval is sensitive to the density of the particles and thus the degree of riming.
The combined application of Eq. (15) in regions where $Z_{DR} > 0.4$ dB and Eq. (17) elsewhere, as suggested by Carlin et al. (2021), leverages the strengths of both formulas and is denoted as $IWC_{Carlin}$ estimator in the following.

Estimating snow concentration $N_t$ is a challenging task, and it is almost impossible to derive it with acceptable accuracy from single-polarization radar measurements because of the wide variety of ice and snow habits and their microphysical properties. A polarimetric retrieval for the logarithm of the total number concentration of ice particles $N_t$ (in L$^{-1}$) following Ryzhkov et al. (2018) and Ryzhkov and Zrnić (2019) is estimated by

$$\log_{10}(N_t(Z_H, Z_{dp}, K_{DP})) = 0.1Z_H - 2\log_{10}\alpha - 1.33,$$  \hspace{1cm} (18a)

with

$$\alpha \approx 0.78 \left(\frac{Z_{dp}}{K_{DP}}\right).$$  \hspace{1cm} (18b)

Again, using the ratio of $Z_{dp}/K_{DP}$ in this retrieval cancels out the effects of orientation and shape. Another relation for $N_t$ included in our accuracy assessment is given by Carlin et al. (2021):

$$\log_{10}(N_t(Z_H, IWC)) = 6.69 + 2\log_{10}(IWC) - 0.1Z_H. \hspace{1cm} (19)$$

Combined with $IWC_{Carlin}$, Eq. (19) is hereafter referred to as $N_t(Z_H, IWC_{Carlin})$. Note that polarimetric retrieval equations (Eqs. 9, 10, and 15–18a) were derived in the Rayleigh approximation assuming that the density $\rho_s$ of Rayleigh scatterers (ice particles, snowflakes) is inversely proportional to $D$, according to the formula following Brandes et al. (2007):

$$\rho_s(D) = \alpha_0 f_{rim}D^{-1} = \alpha_p D^{-1}, \hspace{1cm} (20)$$

where $\rho_s$ is expressed in grams per cubic centimeter, $\alpha_0$ is a constant, and the prefactor $\alpha_p$ varies with the degree of riming $f_{rim}$, which ranges from 1 for unrimed ice to 5 for heavily rimed ice. For larger non-Rayleigh scatterers like graupel or hail, these polarimetric retrieval equations are not valid (Ryzhkov et al., 2020). Reliable retrievals are only obtainable in areas where $Z_{DR}$ and $K_{DP}$ are not very close to zero, which represents a weakness of polarimetric retrievals. In addition, a recent study showed that polarimetric ice microphysical retrievals following Ryzhkov and Zrnić (2019) provide the best results at cold temperatures, i.e., lower than $-10$ to $-15$ °C (Murphy et al., 2020). Around this temperature interval the dendritic growth layer (DGL) is located and $Z_{DR}$ and $K_{DP}$ exhibit pronounced signals. For this reason, our analysis is restricted to temperatures below $-10$ °C. Table 2 summarizes the retrieval equations used in this study.

4 Methodology

RHI scans obtain vertically high-resolved measurements and thus are well suited for microphysical studies and our accuracy assessment. Furthermore, the sequences of successive RHIs performed during the OLYMPEX campaign in azimuthal sectors of 22° provide high-resolution 3D measurements in this predefined region. Our newly introduced RHI sector vertical profile (RSVP) technique provides noise-reduced quasi-vertical profiles of polarimetric variables obtained by azimuthal averaging of RHI sector scans in a convenient height vs. time format. This method was inspired by both the quasi-vertical profile (QVP; Ryzhkov et al., 2016) methodology and the ability of CVPs to follow flight segments of research aircraft, and on top it takes advantage of the RHI scan mode. Figure 1 illustrates the RSVP technique introduced here for the matching with airborne in situ measurements.

To create a RSVP, a series of RHIs is first averaged within the azimuthal range $\alpha$. The range of elevation angles $\theta$ considered can be adapted, e.g., to minimize ground clutter effects.

In the second step, windows of the desired size and position are selected along the range axis of the already azimuthally averaged RHI (marked in red in Fig. 1) and averaged along the chosen range interval as well. Based on the selected and averaged volume, a mean profile is computed representing the vertical columns that are combined and displayed in the RSVP. In this way, RSVPs enable both the
joint analysis with fixed ground-based measurements provided by vertically pointing devices like, e.g., micro rain radars (MRRs) and the tracking of research aircraft within the sector covered by the RHIs for the matching with airborne measurements. In the latter case, the selected volume taken into account in the averaging process changes with time. Additionally, vertical averaging is applied, with 75 m bins to match the aircraft track and account for aircraft altitude fluctuations. Four columns with a length on the range of 4.5 km each starting at 2 km distance from the radar are used for tracking the aircraft. The first 2 km were omitted because of known inconsistencies in the transmitters and reduced polarimetric information content. Also, the maximum range considered in this analysis was 20 km from the radar in order to reduce partial beam blockage by surrounding mountains. The temporal resolution of the RSVP technique depends for sure on the scan schedule. During the OLYMPEX campaign, the 22 RHIs measured in the azimuthal sector were available every 4 min interspersed with a 2 min PPI scan after each two RHI sector scans.

Figure 2 shows as an example RSVPs of \( Z_h, Z_{dr}, \rho_{hv}, \) and \( K_{DP} \) for a complex occluded front observed on 18 December 2015. Similar RSVPs have been generated for all 20 flights during OLYMPEX totalling approximately 60 flight hours (not shown here). The event displayed in Fig. 2 exhibits throughout clearly visible ML signatures in \( Z_h, Z_{dr}, \) and \( \rho_{hv} \), roughly following the temporal evolution of the 0°C isotherm. Within the DGL, located at temperatures between −10 and −15°C, also bands of enhanced \( Z_{dr} \) and \( K_{DP} \) are visible.

For the accuracy assessment, the microphysical retrievals introduced in Sect. 3 are calculated based on the RSVPs and displayed in a similar manner, considering only data above the ML. For this purpose, only radar data with \( Z_{DR} > 0.1 \text{ dB}, Z_H > 0 \text{ dBZ}, K_{DP} > 0.01 \text{ km}^{-1}, \) and \( \rho_{hv} > 0.7 \) are used.

A direct comparison of airborne in situ measurements with ground-based retrievals requires a careful matching in both space and time. Research aircraft measure along flight trajectories, while the RSVP technique uses stationary ground-based radars monitoring a volume at flight altitude. The aircraft measurements of IWC, \( D_m \), and \( N_i \) are averaged along the respective flight path sections and compared to the radar retrievals in the according columns and at the corresponding aircraft altitude. Similarly, the flight altitude and measured environmental parameters (e.g., temperature) are averaged for the time intervals within each column. The temperature information enables us to exclude bright band effects. Herein, the averaged in situ observations are assumed to be characteristic of the entire collocated radar volume. Figure 3 illustrates the matching of microphysical retrievals based on RSVP data with airborne in situ measurements. During OLYMPEX, the UND Citation II research aircraft performed 148 transects over DOW at different altitudes. In our study, only the aircraft measurements between 3 and 7 km height

### Table 2. Table of retrieval equations used in this work (\( D_m \) is in units of mm, IWC is in units of g m\(^{-3} \), and \( N_i \) is in units of L\(^{-1} \); \( T \) is in units of °C.)

<table>
<thead>
<tr>
<th>Retrieval</th>
<th>Formula</th>
<th>Type</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>( D_m(Z_h) )</td>
<td>( 1.45Z_h^{0.25} )</td>
<td>Empirical</td>
<td>Skofronick-Jackson et al. (2019)</td>
</tr>
<tr>
<td>( D_m(Z_h) )</td>
<td>( 1.06Z_h^{0.271} )</td>
<td>Empirical</td>
<td>Matrosov et al. (2019)</td>
</tr>
<tr>
<td>( D_m(Z_{dp}, K_{DP}) )</td>
<td>( -0.1 + 2.0(Z_{dp}/K_{DP})^{1/2} )</td>
<td>Empirical</td>
<td>Ryzhkov and Zrnić (2019)</td>
</tr>
<tr>
<td>( D_m(Z_h, K_{DP}) )</td>
<td>( 0.67(Z_h/K_{DP})^{1/3} )</td>
<td>Empirical</td>
<td>Bukovčić et al. (2020)</td>
</tr>
<tr>
<td>( \log_{10}(IWC(Z_H, T)) )</td>
<td>( 0.06Z_H - 0.0197T - 1.7 )</td>
<td>Empirical</td>
<td>Hogan et al. (2006)</td>
</tr>
<tr>
<td>( \log_{10}(IWC(Z_H, T)) )</td>
<td>( 0.06Z_H - 0.0121T - 1.92 )</td>
<td>Model</td>
<td>Hogan et al. (2006)</td>
</tr>
<tr>
<td>( IWC(K_{DP}) )</td>
<td>( 0.903K_{DP} + 0.319 )</td>
<td>Empirical</td>
<td>Nguyen et al. (2019)</td>
</tr>
<tr>
<td>( IWC(Z_{dr}, K_{DP}) )</td>
<td>( 0.136K_{DP} + 0.037((Z_{dr}^{-1} - 1)^{-1}) )</td>
<td>Empirical</td>
<td>Nguyen et al. (2019)</td>
</tr>
<tr>
<td>( IWC(Z_{dr}, K_{DP}) )</td>
<td>( 4 \times 10^{-2}((Z_{dr}^{-1} - 1)^{-1}) )</td>
<td>Theoretical</td>
<td>Ryzhkov and Zrnić (2019)</td>
</tr>
<tr>
<td>( IWC(Z_h, K_{DP}) )</td>
<td>( 0.31K_{DP}^{0.66}Z_h^{0.28} )</td>
<td>Empirical</td>
<td>Bukovčić et al. (2020)</td>
</tr>
</tbody>
</table>
| \( IWC_{Carlin} \) | \( \begin{cases} IWC_{II}(Z_{dr}, K_{DP}) & \text{if } Z_{DR} > 0.4 \text{ dB}, \\
IWC(Z_h, K_{DP}) & \text{otherwise} \end{cases} \) | Theoretical and empirical | Carlin et al. (2021) |
| \( \log_{10}(N_i(Z_H, Z_{dp}, K_{DP})) \) | \( 0.1Z_H - 2\log_{10}(IWC) - 1.33 \) | Theoretical | Ryzhkov and Zrnić (2019) |
| \( \log_{10}(N_i(Z_H, IWC)) \) | \( 6.69 + 2\log_{10}(IWC) - 0.1Z_H \) | Theoretical and empirical | Carlin et al. (2021) |
are used. The use of RSVP columns allows for the analysis of multiple collocated data points within a single overpass. Here, the selected length of 4.5 km along the range for the RSVP columns ensures that the duration of an aircraft flying at approximately 100 m s\(^{-1}\) is sufficiently long within the column (\(t \geq 30\) s). Flight intercepts of less than 30 s in duration were discarded from the analysis as they may not adequately represent the respective RSVP columns and to ensure statistical reliability. Too short flight intercepts occurred, e.g., when the aircraft shortly flew along the edge of a column and then left the sector again (e.g., yellow column in right panel of Fig. 3) or did not pass the RHI sector at all. Note that the limited temporal resolution of RSVPs (here 4 min) and the comparatively short flight segments can lead to a temporal mismatch between the two datasets of up to 3 min in this study.

5 Accuracy assessment of the ice microphysical retrievals

This section presents the resulting accuracies for all polarimetric and non-polarimetric retrievals outlined in Sect. 3 exploiting the collocated in situ measurements available during the OLYMPEX campaign. The in situ measurements are assumed to be the truth in this study despite similar existing uncertainties, e.g., with the assumed mass–dimension relationship. As statistical measures for the agreement between the in situ measurements and the different radar-derived retrievals, the mean and median retrieved-to-measured ratio (RMR), the mean difference (bias), the root mean square error (RMSE), and Pearson’s correlation coefficient \(r\) are considered. Results for all ice microphysical retrievals introduced in Sect. 3 are shown in Table 3. The retrieval analysis identifies in terms of RMSE \(\text{IWC}_{\text{Carlin}}\), \(N_i(Z_H, \text{IWC}_{\text{Carlin}})\), and \(D_m(Z_{dp}, K_{DP})\) as the best-performing set for the three quantities considered. It can be seen that the use of polarimetry clearly improves the estimates of \(\text{IWC}\) and \(D_m\) compared to the conventional, non-polarimetric retrievals. In particular, \(\text{IWC}_{\text{Carlin}}\) can improve \(r\) by 7\% over \(\text{IWC}(Z_H, T)\) and reduce RMSE by 37\%. An even greater improvement occurs for the estimation of \(D_m\), where the use of \(D_m(Z_{dp}, K_{DP})\) over both non-polarimetric \(D_m\) retrievals increases \(r\) by 15\% compared to both non-polarimetric \(D_m\) retrievals. \(D_m(Z_{dp}, K_{DP})\) also brings a reduction in RMSE of 47\% compared to \(D_m^1(Z_h)\). An analysis of \(\text{IWC}(Z_{dr}, K_{DP})\) in the ice regions of tropical clouds during seven flights of the HAIC-HIWC field campaign in Cayenne, French Guiana, showed an overall correlation between in situ and estimated IWCs of 0.72 and a mean RMSE of 0.52 g m\(^{-3}\) (Nguyen et al., 2019). In our
A. Blanke et al.: Evaluation of polarimetric microphysical retrievals

Figure 3. A schematic illustration from the collocation of RSVP columns with in situ measurements. Each colored column can be created according to the RSVP procedure depicted in Fig. 1. Gray shaded areas represent omitted data very close to the radar. Panel (a) shows the side view and panel (b) the top view of the aircraft flying through the columns.

case, a lower RMSE of 0.26 g m$^{-3}$ can be observed for their IWC($Z_{dr}, K_{DP}$) and a systematically higher correlation of 0.97. Again, the advantages of using polarimetry over non-polarimetry are evident in this retrieval. It is interesting that IWC($K_{DP}$) shows the highest correlation with 0.98, although $Z_{dr}$ is not included. Nevertheless, both polarimetric estimators based on optimal fitting parameters exhibit a higher RMSE compared to all other polarimetric IWC retrievals, manifested in a systematic overestimation. A possible explanation for the overestimation may be that they were optimized for the tropical climate region characterized by an average higher IWC.

For a more detailed look at the evaluation procedure to obtain information about quantitative statistics and accuracies, we consider an example (Fig. 4) for the best-performing IWC retrieval. Out of 20 campaign flights, 10 flight missions were selected and are listed in Table 4 that have good data quality (in situ and radar) and meet all filter criteria. With the thresholds applied, the correlation of IWC$_{Carlin}$ improves from $r = 0.91$ to $r = 0.96$. The slope of the filtered regression is 1.03, which is closer to the 1 to 1 line than the unfiltered green regression line. The important information here is that the filtering via thresholds provides us with a more reliable database for the analysis.

The repeated analysis without a temperature threshold applied showed a significant decrease in the correlations for all retrievals (not shown). Figure 5 shows this best-performing set of ice microphysical retrievals, together with the in situ data. Overall, we see a tendency towards a slight overestimation of IWC at warmer ($T \geq -14 \, ^\circ C$) and an underestimation at colder temperatures ($T \leq -27 \, ^\circ C$), with the largest in situ standard deviation at colder temperatures. IWC$_{Carlin}$ yields a high correlation of $r = 0.96$, the lowest RMSE of 0.19 g m$^{-3}$, and a near-zero bias of $-0.04 \, g \, m^{-3}$. Carlin et al. (2021) found some evidence that using IWC$_{Carlin}$ for the initialization of a 1D spectral bin model results in a more constrained forecast with respect to the snowfall start time compared to using the IWC($Z_{H}, K_{DP}$) retrieval only. The outstanding performance of IWC$_{Carlin}$ within this accuracy assessment is in line with these findings.

The polarimetric retrieval for $N_t(Z_{H}, \text{IWC}_{Carlin})$ poses a greater challenge (Fig. 5, middle panel). It shows a high variability and more pronounced deviations from the in situ measurements but still reaches a convincing correlation of $r = 0.91$. Similar to IWC$_{Carlin}$, $N_t(Z_{H}, \text{IWC}_{Carlin})$ shows an
Table 3. Correlations ($r$), slopes, and intercepts from least-squares fits, root mean square error (RMSE), biases, and mean and median retrieved-to-measured ratio (RMR) for each microphysical property and for all retrievals. Best values of each statistical measure across every microphysical retrieval type are highlighted in boldface font.

<table>
<thead>
<tr>
<th>IWC (g m$^{-3}$)</th>
<th>$r$</th>
<th>Slope</th>
<th>Intercept</th>
<th>RMSE</th>
<th>Bias</th>
<th>RMR mean</th>
<th>RMR median</th>
<th>Publication</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{IWC}^I(Z_H, T)$</td>
<td>0.90</td>
<td>0.88</td>
<td>-0.04</td>
<td>0.30</td>
<td>-0.11</td>
<td>1.23</td>
<td>1.59</td>
<td>Hogan et al. (2006)</td>
</tr>
<tr>
<td>$\text{IWC}^I(Z_H, T)$</td>
<td>0.91</td>
<td>1.40</td>
<td>-0.05</td>
<td>0.30</td>
<td>0.10</td>
<td>0.79</td>
<td>1.00</td>
<td>Met Office Model</td>
</tr>
<tr>
<td>$\text{IWC}_{\text{Comb}}(Z_H, T)$</td>
<td>0.95</td>
<td>1.18</td>
<td>-0.06</td>
<td>0.20</td>
<td>0.02</td>
<td>0.95</td>
<td>1.11</td>
<td>Nguyen et al. (2019)</td>
</tr>
<tr>
<td>$\text{IWC}(K_{DP})$</td>
<td>0.98</td>
<td>1.29</td>
<td>-0.42</td>
<td>0.28</td>
<td>-0.22</td>
<td>1.46</td>
<td>2.02</td>
<td>Nguyen et al. (2019)</td>
</tr>
<tr>
<td>$\text{IWC}^I(Z_{\text{dr}}, K_{DP})$</td>
<td>0.97</td>
<td>1.20</td>
<td>-0.34</td>
<td>0.26</td>
<td>-0.21</td>
<td>1.43</td>
<td>1.97</td>
<td>Ryzhkov et al. (2018)</td>
</tr>
<tr>
<td>$\text{IWC}^I(Z_{\text{dr}}, K_{DP})$</td>
<td>0.94</td>
<td>0.87</td>
<td>-0.02</td>
<td>0.24</td>
<td>-0.10</td>
<td>1.21</td>
<td>1.25</td>
<td>Bukovčić et al. (2018)</td>
</tr>
<tr>
<td>$\text{IWC}^I(Z_{\text{H}}, K_{DP})$</td>
<td>0.94</td>
<td>1.00</td>
<td>-0.10</td>
<td>0.23</td>
<td>-0.10</td>
<td>1.22</td>
<td>1.76</td>
<td>Carlin et al. (2021)</td>
</tr>
</tbody>
</table>

$n\log_{10}$ (L$^{-1}$)

| $n(Z_H, Z_{dp}, K_{DP})$ | 0.88 | 1.13 | -0.33 | 0.46 | -0.18 | 1.17 | 1.33 | Ryzhkov et al. (2018) |
| $n(Z_H, \text{IWC}_{\text{Carlin}})$ | 0.91 | 1.38 | -0.52 | 0.43 | -0.09 | 1.08 | 1.27 | Carlin et al. (2021) |

$D_m$ (mm)

| $D_m^I(Z_H)$ | 0.79 | 0.55 | 0.04 | 2.12 | -1.82 | 1.78 | 1.86 | Skofronick-Jackson et al. (2019) |
| $D_m^I(Z_H)$ | 0.79 | 0.64 | 0.19 | 1.40 | -0.99 | 1.42 | 1.48 | Matrosov et al. (2019) |
| $D_m(Z_H, K_{DP})$ | 0.94 | 2.60 | -0.85 | 1.38 | 1.10 | 0.53 | 0.54 | Bukovčić et al. (2018) |
| $D_m(Z_{dp}, K_{DP})$ | 0.91 | 1.59 | 0.01 | 1.13 | 0.87 | 0.63 | 0.74 | Ryzhkov et al. (2018) |

Figure 5. Collocated aircraft in situ data in chronological order (colored dots) and the best-performing set of ice microphysical retrievals based on RSVP data (solid blue lines) for 10 flight missions. Plots represent from top to bottom IWC$^I(Z_H, \text{IWC}_{\text{Carlin}})$, $n(Z_H, \text{IWC}_{\text{Carlin}})$, and $D_m(Z_{dp}, K_{DP})$. Shadings show $\pm 1\sigma$ (gray) and standard error of the mean (blue) calculated via Gaussian error propagation. Colors following the color bar indicate the respective temperatures (in °C). Vertical bars represent in situ standard deviations.
overestimation at warmer \(T \gtrsim -14 \, ^\circ \text{C}\) and underestimation at colder temperatures \(T \lesssim -27 \, ^\circ \text{C}\). Pronounced \(N_t\) deviations at warmer temperatures are visible for the collocated data points 29 and 31, with the former also revealing large standard deviations on the in situ side. The outlier at data point 9 exhibits \(K_{DP}\) values below 0.1 \, km^{-1}, which may indicate deficiencies of this retrieval for very low \(K_{DP}\) values, but the in situ values show relatively high variability as well. The \(D_m(Z_{dp}, K_{DP})\) retrieval (Fig. 5, bottom panel) estimates particle size especially well at colder temperatures.

Mixtures of aggregates with different sizes coexisting near the ML (solid region) may explain the high in situ standard deviations at warmer temperatures (e.g., data point 29 where \(D_m\) is in excess of 5 mm). Furthermore, the occurrence of an increased number of aggregates reducing the information content of \(K_{DP}\) can result in an underestimation of \(D_m(Z_{dp}, K_{DP})\), which makes \(D_m\) estimation via polarimetric retrievals close to the freezing level challenging. In line with our analyses, Murphy et al. (2020) indicated larger errors with polarimetric \(N_t\) and \(D_m\) retrievals in the immediate vicinity of the ML. Despite this underestimation, a convincing correlation of \(r = 0.91\) is obtained.

Direct comparisons of both IWC\((Z, T)\) retrievals with the IWC\(_{\text{Carlin}}\) retrieval reveals that both non-polarimetric retrievals show a worse performance in terms of lower correlation and higher RMSE (see Table 3). Note that the use of temperature information from soundings or models would further reduce the performance of IWC\((Z, T)\) retrievals due to increased uncertainties compared to in situ temperature data recorded during the flights. Figure 6 indicates that each IWC\((Z, T)\) shows a better performance in a certain temperature range. Accordingly, we combined them using Eq. (11) for \(T \leq 15 \, ^\circ \text{C}\) and Eq. (12) elsewhere; i.e., the DGL located between approximately \(-10\) to \(-15 \, ^\circ \text{C}\) providing the optimal conditions for the depositional growth of ice serves here as the boundary for the two retrievals. This combination, hereafter referred to as IWC\(_{\text{Comb}}\)(\(Z_H, T\)), shows promise in estimating IWC with non-polarimetric data more precisely, as it reduces the RMSE to 0.20 and increases the correlation to 0.95 (Table 3). The bias of the combined method with a value of 0.02 g m\(^{-3}\) is slightly closer to 0 compared to IWC\(_{\text{Carlin}}\). However, a more extensive dataset is needed to corroborate this finding.

Heymsfield et al. (2008) also evaluated non-polarimetric IWC retrieval methods, including the IWC\(^4\)(\(Z_H, T\)) variant for 95 GHz, using test datasets derived from in situ microphysical measurements. The retrieved-to-measured ratio (RMR) was calculated via dividing the mean or median of the retrieved quantity by the measured quantity, with a selected range of \(0.75 < \text{RMR} < 1.25\) indicating “good” agreement between retrievals and measurements. Their IWC\(^4\)(\(Z_H, T\)) showed a tendency to underestimate IWC at low temperatures and overestimate it at warm temperatures, consistent with our results. Overall, their non-polarimetric radar-only approach and radar temperature retrievals yielded a mean (median) RMR of 1.29 (1.20). Our analysis shows a similar mean RMR of 1.23 for IWC\(^4\)(\(Z_H, T\)) but a slightly higher median RMR value of 1.59. It is noteworthy that the polarimetric IWC\(_{\text{Carlin}}\) outperforms IWC\(^4\)(\(Z_H, T\)) and their radar-only and radar temperature retrievals, with a mean RMR close to unity (see Table 3).

Figure 7 also demonstrates the pronounced biases associated with the \(D_m^4(Z_h)\) and \(D_m^m(Z_h)\) retrievals, with especially large deviations (strong overestimation) at colder temperatures. This result underlines again the importance of the key variable \(K_{DP}\) in ice and reveals the notorious inaccuracies of conventional \(D_m\) retrievals. However, these power laws could potentially be used in combination with polarimetric \(D_m\) retrieval in areas near the ML in a hybrid fashion to obtain a more accurate estimate of \(D_m\).

Our analysis of \(N_t(Z_H, \text{IWC}_{\text{Carlin}})\) shows good agreement in terms of mean RMR (1.08), whereas none of the \(D_m\) retrievals in this evaluation fall within the chosen RMR range, with the best-performing \(D_m(Z_{dp}, K_{DP})\) exhibiting a RMR median value close to the lower limit with a value of 0.74. Overall, our analysis is consistent with findings by Murphy et al. (2020), who found a strong underestimation of the \(D_m(Z_{dp}, K_{DP})\) retrieval near the ML. Similarly, our analysis shows the best results in regions with high \(Z_{DR}\) and \(K_{DP}\), such as in the DGL, and the worst just above the freezing level, where \(Z_{DR}\) and \(K_{DP}\) signatures nearly vanish as a result of aggregation processes (Ryzhkov et al., 1998).

6 Discrepancy analyses

For a more detailed analysis and improved understanding, examples of the most pronounced discrepancies encountered between in situ and retrieved quantities are spotlighted and presented, together with CPI imagery, HVPS samples, and/or RSVPs.

On 10 December 2015 at 15:58 UTC (data point 22, see top panel in Fig. 5), the aircraft entered the RSVP column at a distance of 15.5 to 20 km from the DOW. In this flight seg-

Table 4. List of selected Citation II flight missions after applying filter criteria with resulting collocated data points.

<table>
<thead>
<tr>
<th>Dates</th>
<th>Flight times (UTC; to nearest 15 min)</th>
<th>Data points</th>
</tr>
</thead>
<tbody>
<tr>
<td>12 November</td>
<td>19:30–22:30</td>
<td>0–1</td>
</tr>
<tr>
<td>13 November</td>
<td>15:00–17:45</td>
<td>2–3</td>
</tr>
<tr>
<td>1 December</td>
<td>22:45–01:45</td>
<td>4–8</td>
</tr>
<tr>
<td>4 December</td>
<td>13:00–16:00</td>
<td>9–11</td>
</tr>
<tr>
<td>5 December</td>
<td>14:45–18:00</td>
<td>12–14</td>
</tr>
<tr>
<td>10 December</td>
<td>14:45–17:00</td>
<td>15–22</td>
</tr>
<tr>
<td>12 December</td>
<td>17:00–20:15</td>
<td>23–26</td>
</tr>
<tr>
<td>13 December</td>
<td>15:45–19:15; 20:00–23:15</td>
<td>27–34</td>
</tr>
<tr>
<td>18 December</td>
<td>01:15–04:30</td>
<td>35–37</td>
</tr>
</tbody>
</table>
Figure 6. Collocated aircraft in situ data in chronological order (colored dots), the IWC\textsuperscript{I}(Z, T) (green), the IWC\textsuperscript{II}(Z, T) (gray), and the IWC\textsubscript{Carlin} (blue) of RSVP data (solid lines) for 10 flight missions. Vertical bars have the same meaning as in Fig. 5.

Figure 7. Collocated aircraft in situ data in chronological order (colored dots), the $D_m$ power laws (yellow and green), and the $D_m(Z_{dp}, K_{DP})$ retrieval (gray) of RSVP data (solid lines) for 10 flight missions. Vertical bars have the same meaning as in Fig. 5.

The most pronounced discrepancy in terms of $D_m$ emerges on 13 December 2015 at 16:17 UTC (data point 29, see lower panel in Fig. 5) when the aircraft entered the RPI column at a distance of 6.5 to 11 km from the DOW. Here, the Citation II measured a mean temperature of $-13$ °C at a mean altitude of 2.95 km. Retrieved $D_m(Z_{dp}, K_{DP})$ strongly underestimates the in situ observed $D_m$ peak exceeding 5 mm, with an underestimation of almost 3 mm. Note that in this case there is no temporal mismatch. Figure 9 illustrates sampled particles captured along the relatively narrow flight path of the aircraft. It seems the smaller in situ sample is dominated by larger particles, while the larger RSVP volume also accounts for the immediate vicinity where more smaller particles may be present. These possibly affect $D_m(Z_{dp}, K_{DP})$ significantly due to averaging and may not properly resolve the regions with enhanced in situ $D_m$ peaks. In addition, the DOW sector scan started at 16:14 UTC, and thus the RSVP volume also contains particles not monitored by the aircraft entering the sector later. The particles may be smaller and could explain the discrepancy between in situ observed $D_m$ and retrieved $D_m(Z_{dp}, K_{DP})$. It is also worth mentioning that...
Figure 8. RSVPs of IWC_{Carlin} at ranges between 15.5 and 20 km from the DOW on 10 December 2015 between 14:30 and 16:54 UTC. Overlaid dashed lines display the −5, −10, −15, and −20 °C isotherms from ERA5 at the DOW location. The collocated data point is indicated by the black cross.

Figure 9. HVPS sample images of hydrometeor shadows on 13 December 2015 at 16:18 UTC. The images correspond to the trajectory within the DOW sector scan starting at 16:17 UTC. The height of each panel represents 19.2 mm.

In agreement with the RICE measurements, indicating ongoing riming in the flight segment under discussion. Moreover, a clear sagging of the ML in terms of Z_{DR} and ρ_{hv} (not shown) in the RSVPs supports the riming hypothesis, since rimed ice particles fall with increased velocities and therefore melt at lower altitudes. Such a sagging signature in radar images can be associated with riming processes (Kumjian et al., 2016), and since riming, most likely present in this case, results in both higher ϕ and f_{rim} values than assumed in the derivation of the D_{m}(Z_{dp}, K_{DP}) retrieval, the observed underestimation is in line with the known shortcomings of the retrieval for such conditions. Additionally, for riming conditions present, the mass–dimension relation assumed for in situ measured IWC is also not valid anymore. In general, the precise effect of riming on mass–dimension relations is poorly understood (Tridon et al., 2019), but these processes considerably modify inherent parameters. With an increasing degree of riming, higher values of the prefactor a and the exponent b in Eq. (6) are expected. The latter can reach a maximum b close to 3 (sphere-like geometry) when graupel-like particles filled with rime are present. Thus, the observed overestimation of IWC_{Carlin} could at least partly be explained by assumed inappropriate parameters in the mass–dimension relationship following Brown and Francis (1995). Using a relation that accounts for more knowledge of ice particle masses (e.g., Heymsfield et al., 2010) that is valid for higher degrees of riming (e.g., Leinonen and Szyrner, 2015) or that uses multiple mass–dimension relations for different D_{mm} ranges, as proposed in Ding et al. (2020), may reduce deviations between in situ and retrieved IWC in this case. Only directly measured in situ IWC can provide IWC without the need for any assumptions. However, such measurements were not available during the OLYMPEX campaign (Tridon et al., 2019).

7 Conclusions

Data collected during the OLYMPEX campaign (Houze et al., 2017) conducted in late 2015 provided a comprehensive database including ground-based polarimetric X-band radar measurements and airborne in situ cloud measure-
data assimilation. The key results of the study are as follows.

1. State-of-the-art microphysical retrievals exploiting polarimetric radar measurements to estimate IWC (RMSE = 0.19 g m$^{-3}$), $N_l$ (RMSE = 0.43 L$^{-1}$), and $D_m$ (RMSE = 1.13 mm) achieve quite high agreement with airborne in situ measurements, especially at cold temperatures.

2. Overall, polarimetric retrievals are superior to conventional Z-based retrievals, but combinations with non-polarimetric retrievals have potential to improve deficiencies directly above the ML.

3. The hybrid polarimetric IWC$_{Carlin}$ retrieval outperforms all other IWC estimates in terms of RMSE and shows a high $r$ of 0.96.

4. Compared to IWC retrievals, the $N_l$ and $D_m$ retrievals show larger uncertainties and should be further improved in the future.

This study clearly demonstrates the added value of multiparameter retrievals as proposed by Carlin et al. (2021) for the IWC. Combining the strengths of retrievals can be applied to other ice microphysical properties and ultimately also hints at the potential for developing combinations of polarimetric and non-polarimetric retrievals. Further independent evaluation studies particularly focusing on such hybrid retrievals are required.

Future work using an additional particle classifier (e.g., Praz et al., 2018; Przybylo et al., 2022) capable of identifying and classifying each particle along flight transects with high-resolution image data sampled by particle imagers will provide an even more in-depth evaluation. Specifically, extended classifiers that possibly allow riming degree estimation, as proposed in Przybylo et al. (2022), may provide an avenue for studying riming cases and relating these observations to retrieval assumptions. These classifiers hold the potential to be utilized for the refinement and development of future ice microphysical retrieval methods.

The exploitation of the OLYMPEX data represents another piece of the mosaic towards a comprehensive evaluation of (polarimetric) microphysical retrievals. While previous evaluation studies focused on C- and S-band radar data, this study also emphasizes the added value of X-band radars for the exploitation of microphysical retrievals and related process studies. However, in light of potential applications in model evaluation and data assimilation, C- and S-band radars are of great interest because of their national operational availability. Recently, even more field campaigns have been exploited or their analysis is currently underway. Examples are the Federal-Aviation-Administration-led In-Cloud ICing and Large-drop Experiment (ICICLE; Bernstein et al., 2021) and the NASA-led Investigation of Microphysics and Precipitation for Atlantic Coast-Threatening Snowstorms (IMPACTS; McMurdie et al., 2019). Two flight legs of a winter storm case from the latter campaign have already been used by Dunnavan et al. (2022) for a radar retrieval evaluation. With respect to previous experimental evaluation studies, the more convincing accuracy of the best-performing polarimetric retrievals identified in this study gives us further confidence in their application. The benefit of using a shorter wavelength than that of S-band radars, as suggested in Ryzhkov et al. (1998) for the verification of their proposed IWC$_4$(Z$_{dr}$, $K_{DP}$), indeed exhibited a noticeably lower RMSE in our study (0.24 g m$^{-3}$ vs. 0.4 g m$^{-3}$). Using the same IWC$_4$(Z$_{dr}$, $K_{DP}$) retrieval as Nguyen et al. (2019), we were further able to achieve half the RMSE and a systematically higher correlation with our method based on RHI sector scans, even though both studies used an X-band radar. More recent retrievals were able to achieve even smaller RMSEs. The almost consistent underestimation of $D_m$ (Z$_{dp}$, $K_{DP}$) shown in Murphy...
et al. (2020) is in line with our results. However, in contrast to their study, we could not only attribute the deficits to warmer temperature regimes but also demonstrate the accurate estimation at cold temperatures. As a result, an important open question for future research concerns the deficient performance of retrievals directly above the ML and requires new approaches to obtain accurate estimators in this region.

**Data availability.** OLYMPEX data used in this study were obtained from the NASA EOSDIS GHRC OLYMPEX data archive (https://doi.org/10.5067/GPMGV/OLYMPEX/DATA101, Petersen et al., 2018, and https://doi.org/10.5067/GPMGV/OLYMPEX/DOW/DATA201, Houze et al., 2018). The ERA5 data are stored at the Climate Data Store from the ECMWF and are accessible via https://doi.org/10.24381/cds.bd0915c6 (Hersbach et al., 2023).

**Author contributions.** AB and ST jointly developed the concept and methodology for this work. Data handling and analysis were performed by AB with contributions from MM. Visualization was carried out by AB, who also led the writing with input from ST. AJH processed the Citation II data collected during OLYMPEX and provided scientific expertise on in situ data. All authors contributed to the proofreading and added valuable suggestions to the final draft.

**Competing interests.** The contact author has declared that none of the authors has any competing interests.

**Disclaimer.** Publisher’s note: Copernicus Publications remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Special issue statement.** This article is part of the special issue “Fusion of radar polarimetry and numerical atmospheric modelling towards an improved understanding of cloud and precipitation processes (ACP/AMT/GMD inter-journal SI)”. It is not associated with a conference.

**Acknowledgements.** The research was carried out in the framework of the Priority Programme SPP-2115 “Polarimetric Radar Observations meet Atmospheric Modelling (PROM)” funded by the German Research Foundation (DFG). We also thank all the participants of OLYMPEX for collecting the data used in this study. Similarly, we acknowledge the support of Kai Mühlbauer and the open-source radar library wradlib (https://docs.wradlib.org/en/stable/index.html, last access: 11 November 2022) regarding the processing of radar data.

**Financial support.** This research has been supported by the Deutsche Forschungsgemeinschaft (grant nos. TR 1023/13-1 and VO 1504/5-1) and the National Aeronautics and Space Administration (grant no. 80NSSC20K0897 to Andrew J. Heymsfield).

This open-access publication was funded by the University of Bonn.

**Review statement.** This paper was edited by Patric Seifert and reviewed by two anonymous referees.

**References**


Kedzuf, N. J., Chiu, J. C., Chandrasekar, V., Biswas, S., Joshi, S., Su, Y., van Leeuwen, P. J., Westbrook, C., Blanchard, Y., and


