

The spatial distribution of convective precipitation - an evaluation of cloud microphysics schemes with polarimetric radar observations

Gregor Köcher¹ and Tobias Zinner¹

¹Meteorologisches Institut, Ludwig-Maximilians-Universität München, Munich, Germany

Correspondence: Gregor Köcher (gregor.koecher@physik.uni-muenchen.de)

Abstract. The representation of cloud microphysics in numerical weather prediction models contributes significantly to the uncertainty of weather forecasts. Polarimetric radar observations are increasingly used to evaluate numerical weather simulations, due to their sensitivity to microphysical properties. Typically, this evaluation is performed for individual case studies, which limits the generalizability of the results. This is particularly problematic for convective precipitation events, which are characterized by high variability due to their small scale and rapid error growth of initial uncertainties, and are often associated with severe weather phenomena. In this study, the performance of microphysics schemes in the simulation of convective precipitation events is evaluated statistically over a 30-day dataset. The aim is to assess the distribution of precipitation into convective and stratiform regions, and the microphysical properties in these regions based on radar reflectivity and differential reflectivity. Within a Weather Research and Forecasting model setup, 5 different microphysics schemes of varying complexity are evaluated. The choice of microphysics schemes has a significant impact on the distribution of precipitation; the median convective area fraction varies by an order of magnitude between the microphysics schemes. These differences are attributed to differences in rain drop size distributions. In the convective core, the FSBM and Morrison schemes frequently lack large rain drops, while the Thompson and P3 schemes simulate too many. Statistical evaluations are important to address the prevailing uncertainty surrounding cloud microphysics. The framework presented in this study can serve as a guide for future statistical evaluations of weather models with polarimetric radar observations.

Copyright statement.

1 Introduction

Simulation of convective precipitation events presents a particular challenge, owing to their small scale and rapid error growth of initial uncertainties (Selz and Craig, 2015; Hohenegger and Schär, 2007). These events are characterized by strong vertical air motions, and a high spatial variability (Powell et al., 2016) and frequently coincide with severe weather phenomena, including flooding, hail, strong winds, and tornadoes (Doswell III, 2001). Despite of their potential for hazardousness, current numerical weather prediction models struggle to accurately predict convective precipitation events. Xue et al. (2017) for example demonstrate that 3 different microphysics schemes produce a substantially different dynamical and thermodynamical

structure of a squall line over the continental United States. These events often consist of a leading convective core and a trailing stratiform region (Houze, 1994), with the convective core characterized by strong updrafts and downdrafts, intense precipitation and winds, and possibly hail. The stratiform region typically has a lower precipitation intensity, but can potentially cover larger areas. Due to the different nature and hazard of these regions, the correct representation of the distribution of precipitation into these regions is important for weather forecasting and hydrology.

The accurate simulation of the convective core and surrounding stratiform precipitation is a significant challenge for cloud microphysics schemes. For example, Shrestha et al. (2022) found that their simulations of three convective storms produced too small convective areas compared to stratiform areas. Han et al. (2019) found that their simulations of a squall line case underestimated total stratiform precipitation, primarily due to the underestimation of precipitation area. Qian et al. (2018) found that the trailing stratiform region of a simulated squall line case was strongly sensitive to the choice of cloud microphysics schemes. Not only the distribution of precipitation into convective core and trailing stratiform region is challenging, but also the representation of microphysical properties in each category. Wu et al. (2021) show that the microphysical characteristics of precipitation within a typhoon differ considerably to radar observations and are mainly determined by the microphysics scheme used. In Köcher et al. (2022), we demonstrate that some of the tested microphysics schemes produce graupel particles that are too large or too dense within convective cores. In a related study, Planche et al. (2019) examined the discrepancies between radar observations and simulations for a squall line case, highlighting differences in the vertical profile of rain size distributions within the stratiform region. These discrepancies were primarily attributed to the process of drop breakup. Sun et al. (2023) demonstrate that two evaluated microphysics schemes encounter difficulties in consistently reproducing radar signals in both, the stratiform and convective regions, as well as throughout the vertical profile. These issues are attributed to deficiencies in particle size distributions, density assumptions, and the simulation of ice processes. All of these studies underscore the prevailing uncertainty surrounding the representation of cloud microphysics in numerical weather prediction models, particularly in the context of convective precipitation events.

Cloud microphysical processes include the formation of cloud droplets, ice crystals, and precipitation (Lamb, 2003; Morrison et al., 2020). A characteristic feature of microphysical processes is its inherent complexity, involving numerous (nonlinear) interactions among diverse species of particles and the surrounding air, as well as a broad spectrum of particle shapes and sizes. Given the vast number of particles in a cloud and the complexity of the processes, direct simulation of these processes over the entire range of scales is computationally unfeasible. Consequently, these processes are typically parameterized, meaning their effects on key variables are described without explicitly calculating the processes themselves. This is typically done by so-called cloud microphysics schemes (e.g., Thompson et al., 2008; Morrison and Milbrandt, 2015; Shpund et al., 2019; Seifert and Beheng, 2005). However, uncertainties in microphysical process rates persist due to fundamental gaps in our understanding of the processes, as well as the lack of observations to constrain the schemes (Morrison et al., 2020). Constraining microphysical processes is challenging, as they occur on small scales. Observational constraints can be derived from laboratory studies (e.g., Seidel et al., 2024), in situ observations (e.g., Billault-Roux et al., 2023), and remote sensing observations (e.g., Tana et al., 2023). However, all of these approaches are limited: laboratory studies are typically limited to a small number of particles and idealized setups. In situ observations, for example with aircraft, require much effort and are spatially limited. Remote

sensing observations, such as radar observations, offer much better spatial coverage. However, they generally cannot measure
60 the microphysical properties directly. Instead, these properties are inferred from the observed signals. These limitations hinder
the validation microphysics schemes to improve their representation in models.

In recent years, polarimetric radar observations have been established. For example, in 2015, the German Meteorological
Service (Deutscher Wetterdienst, DWD) finished upgrading its radar network to polarimetric radars (Trömel et al., 2021). Po-
larimetric radars transmit a radar signal at two orthogonal polarizations, and measure the power and phase of the backscattered
65 signals at both polarizations. The received polarimetric signals are affected by numerous particle properties, including shape,
size distribution, orientation, phase, and number concentration of particles within the measurement volume. Consequently,
these polarimetric fingerprints can be utilized to deduce information concerning cloud microphysical processes that affect
these properties. This, in turn, can be used to constrain cloud microphysics schemes in NWP models (Kumjian, 2012; Ryzhkov
et al., 2020).

70 Meanwhile, there are several studies that have employed polarimetric radar observations to evaluate cloud microphysics
schemes in NWP models (e.g., Chen and Liu, 2024; Li et al., 2023; Chen et al., 2021; You et al., 2020; Putnam et al., 2016;
Brown et al., 2016; Wu et al., 2021, and many more). However, all of these studies are limited to case studies. While case
studies are instrumental to understand the processes in detail, they are limited in their generalizability. This limitation is par-
ticular problematic in the context of for convective precipitation, due to large variability between different events (Flack et al.,
75 2019). Only few studies exist that evaluate cloud microphysics schemes statistically over a large dataset. Caine et al. (2013)
statistically compared convective cell properties; however, their statistics were limited to less than 5 days. Stein et al. (2015)
evaluated simulated convective storms of 40 cases, but only for one microphysics scheme and without the use of polarimetric
radar observations. In Köcher et al. (2022), we provided a setup for statistical evaluation of weather model simulations with
polarimetric radar measurements which was applied in Köcher et al. (2023) to statistically evaluate the performance of 5 micro-
80 physics schemes in predicting high-impact weather events, such as hail and heavy rain. Statistical analysis is not only important
for representative results, but also required for uncertainty estimation. As highlighted by Morrison et al. (2020) in one of their
main conclusions, the critical evaluation of model performance necessitates statistically robust remote-sensing approaches.

The objective of this study is to statistically evaluate the performance of 5 cloud microphysics schemes in simulating the
distribution of precipitation into convective and stratiform regions. The aim is to assess the microphysical properties in these
85 regions based on the observed and simulated polarimetric radar signals. Within the scope of this study, we solely focus on the
differential reflectivity (Z_{dr}) and radar reflectivity (Z). Other polarimetric variables, such as specific differential phase (K_{dp})
and correlation coefficient (RHO_{hv}) could provide further insight into the microphysical properties, but will be left for future
studies.

The remainder of this paper is structured as follows: Section 2 describes the data and methods used in this study. Section
90 3 describes the results of the analysis, including the vertical distribution of reflectivity signals (Sect. 3.1), the distribution of
precipitation into convective and stratiform regions (Sect. 3.2), the microphysical properties of the precipitation within these
regions based on polarimetric radar statistics (Sect. 3.3), and the impact of the simulated rain drop size distributions on the
polarimetric radar signals (Sect. 3.4). Section 4 provides a summary and conclusions of the study.

Table 1. Abbreviations, WRF-IDs, and corresponding publications of the employed microphysics schemes.

Name	WRF-ID	Publication
Thompson 2-mom	8	Thompson et al. (2008)
Morrison 2-mom	10	Morrison et al. (2009)
Thompson aerosol-aware	28	Thompson and Eidhammer (2014)
Fast spectral bin (FSBM)	30	Shpund et al. (2019)
Predicted Particle Properties (P3)	50	Morrison and Milbrandt (2015)

2 Data and methods

95 The data and methodologies employed in this study build upon Köcher et al. (2022). The dataset consists in total of 30 days
of precipitation events in the Munich region in South Eastern Germany in 2019 and 2020. These events were observed from
the German Meteorological Service (Deutscher Wetterdienst, DWD) radar network, and simulated with the Weather Research
and Forecasting (WRF) model (Skamarock et al., 2019). A complete list of the measurement days is provided in Köcher et al.
(2022), in their Table A1. Below, we provide a brief overview of the dataset. For a detailed description, we refer to Köcher
100 et al. (2022).

The observational data utilized in this study was provided by the DWD, which operates a network of fully dual polarimetric
C-band radars. These radars provide the polarimetric variables differential reflectivity (Zdr), specific differential phase (Kdp),
and correlation coefficient (RHO_{hv}), in addition to the radar reflectivity (Z). In this study, we solely focus on raw Z and
Zdr data, i.e., without any attenuation correction applied. The network consists of 17 radars that cover the entire area of
105 Germany. For the purpose of this study, the closest radar to Munich, i.e., the radar in Isen, southern Germany was selected. The
operational measurement strategy of the DWD radars involve a volume scan every five minutes, which consist of Plan Position
Indicator (PPI) scans at 11 elevation angles spanning from 0.5 to 25 degrees. Further details on the DWD radar network and
the measurement strategy can be found in Helmert et al. (2014).

The simulations were performed using a regional limited area model setup of WRF, version 4.2. The inner nest is centered
110 over Munich with a horizontal grid spacing of 400 m, 40 vertical levels, and a horizontal extent of 144 km x 144 km. 5 mi-
crophysics schemes were are employed and listed in Table (1). The simulations were initialized with reanalysis data from the
Global Forecast System (GFS; National Centers For Environmental Prediction/National Weather Service/NOAA/U.S. Depart-
ment Of Commerce, 2015) at 0.25 degree horizontal resolution.

In this study, observed polarimetric radar signals are compared to simulated radar signals. The Cloud-resolving model Radar
115 SIMulator (CR-SIM, version 3.33; Oue et al., 2020) is utilized to simulate the radar signals from the WRF model output. CR-
SIM is based on the T-matrix method (Barber and Yeh, 1975) and is capable of providing the polarimetric variables Zdr, Kdp,
and RHO_{hv} based on assumed particle shapes and particle orientations, depending on the simulated hydrometeor classes. The
simulated particle densities and particle size distributions are consistent with the microphysics schemes employed. CR-SIM
does not implement a melting scheme; instead, particles are either completely frozen or completely melted, in accordance with

120 the microphysics schemes. This makes it impossible to simulate melting layer effects such as the “bright band” (Austin and Bemis, 1950), and hence the analysis in this study is focused on heights above and below the melting layer. Furthermore, the attenuations effects are simulated by CR-SIM and applied to the simulated reflectivity (Z) and differential reflectivity (Z_{dr}).

The radar data resolution depends on the distance to the radar. To ensure comparability of the datasets, both are interpolated to a joint Cartesian grid with a horizontal grid spacing of 400 m. The radar data is interpolated to the model grid using
125 an inverse distance interpolation. To ensure a fair comparison between model and radar geometry, as well as to apply the attenuation correction, the model data is first interpolated to a radar grid. This is achieved by considering the radar beam geometry, whereby each pixel is weighted based on the distance to the radar beam center. To allow the application of a cell-tracking algorithm that operates on a Cartesian grid (see Sect. 2.1), the resulting data is then interpolated back to the Cartesian grid using the same inverse distance interpolation as for the radar data.

130 **2.1 Cell-tracking**

An automatic cell-tracking algorithm (tobac; Sokolowsky et al., 2024) is applied to the matched radar and simulation data alike to identify convective cells and their surrounding stratiform regions. Note that this approach is different from the methodology employed in our prior study (Köcher et al., 2022). In the preceding study, the focus was exclusively on convective cores. In contrast, the present study aims for a separation into convective and stratiform regions. Tobac not only tracks convective cores,
135 but also provides the functionality to assign surrounding stratiform regions to the convective cores through the implementation of a segmentation processes based on a watershedding algorithm. The tobac algorithm consists of three main steps:

1. Identification of convective cells based on reflectivity thresholds. For the purpose of this study, a reflectivity threshold of 35 dBZ is used. The location of the maximum reflectivity within a cell is designated as the cell’s center.
2. Segmentation of surrounding stratiform precipitation. A surrounding volume is associated with each convective cell
140 based on watershedding with a fixed threshold of 5 dBZ and 35 dBZ for the stratiform and convective parts, respectively.
3. Tracking of cells over time. Identified cells are linked over time based on the cell position using the python library trackpy (Allan et al., 2021).

Based on the identified convective cells and the associated stratiform regions, the convective area fraction (CAF) is calculated. The CAF is defined as the ratio of the area covered by convective precipitation to the total area covered by precipitation
145 and will be used to describe the distribution of precipitation into convective and stratiform regions. All in all, a total of 8395 convective cells were identified by the tobac algorithm over the 30-day dataset in the radar data while the number of cells in the simulations varied depending on the microphysics scheme and is discussed below. An example of the cell tracking output is shown in Figure 1. For a more detailed description of the radar data, simulation setup, and forward operator, please refer to Köcher et al. (2022). An example case demonstrating the simulation output of the five different microphysics schemes within
150 the NWP model setup is presented in Köcher et al. (2023).

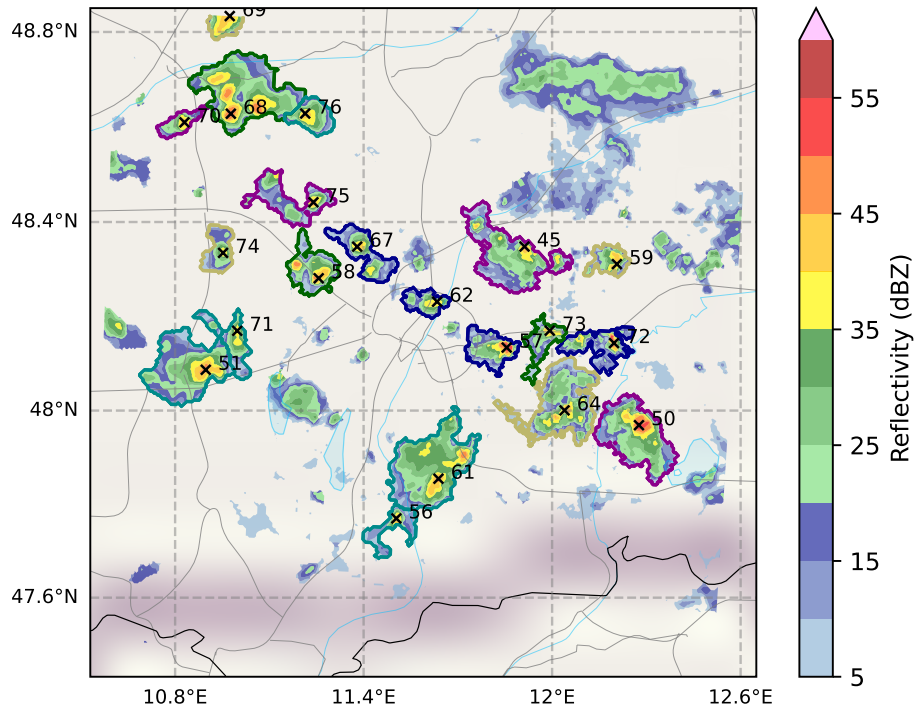


Figure 1. Example tobac tracking applied to DWD radar data on the 20th of June 2020, 10:05 UTC. In color the measured reflectivity. Black 'x' symbols are the identified cells and the black numbers are the corresponding cell IDs. The colored lines around the precipitation areas are the identified stratiform regions that were assigned to the convective cells. Background map data by OpenStreetMap (<http://openstreetmap.org>, last access: April 7, 2025; © OpenStreetMap contributors. Distributed under the Open Data Commons Open Database License (ODbL) v1.0.). Roads, rivers, and lakes made with Natural Earth (<https://www.naturalearthdata.com>, last access: April 7, 2025).

3 Results

In general, the majority of the weather model simulations yielded comparable total surface precipitation amounts, ranging from 295 mm to 321 mm precipitation summed over the entire 30 days over the whole domain (Figure 2). Only the FSBM scheme yielded slightly higher total precipitation amounts, reaching 390 mm. However, while the total precipitation varies by a factor of only around 10% to 30%, the number of identified cells varies by a factor of 2, ranging from 5261 to 11576 across the microphysics schemes. This indicates differences in distribution of precipitation into convective and stratiform regions between the microphysics schemes. For example, FSBM precipitates more water than P3, but shows a much lower number of intense convective cells, indicating more wide spread and less convectively active rain weather in FSBM.

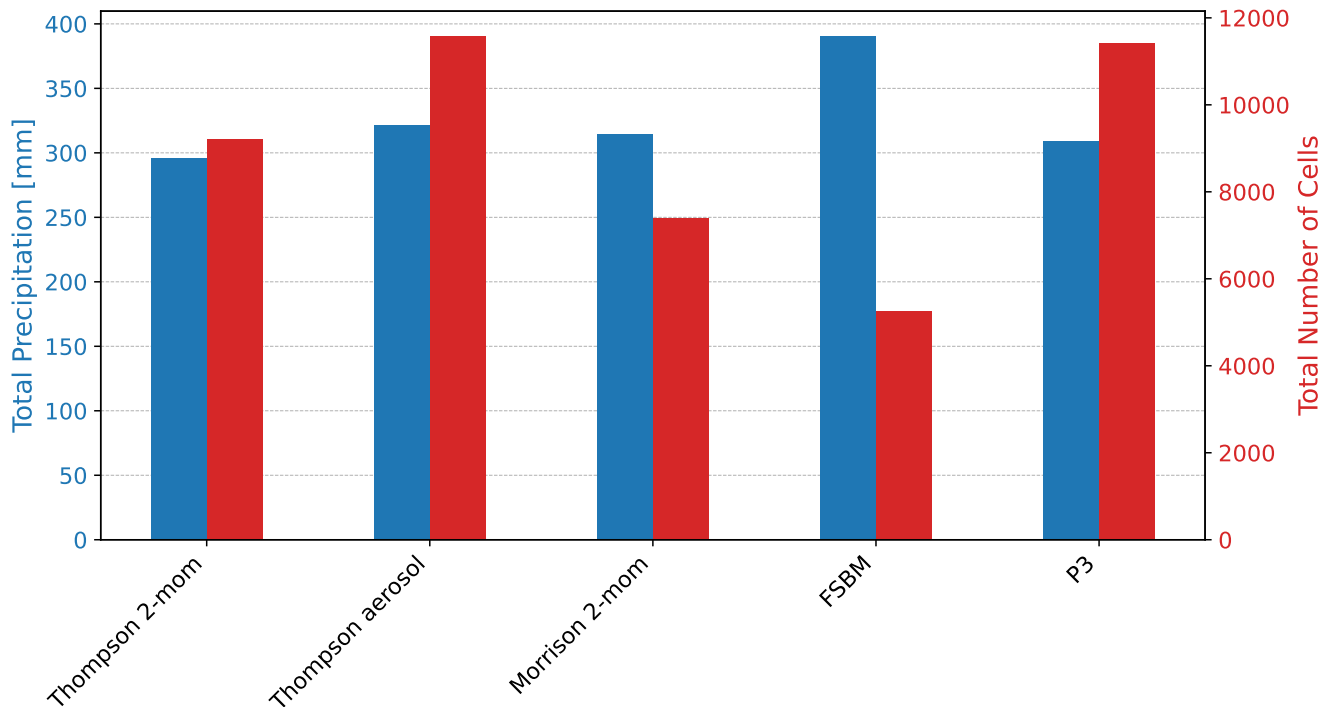


Figure 2. Total WRF simulated precipitation (blue) over all timesteps and all grid boxes as well as total number of identified convective cells (red) for each of the 5 microphysics schemes.

3.1 Vertical distribution of reflectivity signals

160 In order to further evaluate the spatial distribution of precipitation, we next analyze the vertical distribution of the reflectivity signals. Figure 3 shows the vertical distribution of the precipitation fraction based on reflectivity, i.e., the fraction of pixels that exceeded 5 dBZ and 35 dBZ depending on height. A stark difference can be seen between the microphysics schemes up to about 4 km for the convective (35 dBZ) threshold and over the full vertical profile for the (stratiform) 5 dBZ threshold. Most apparent is the Morrison scheme (green) which produces the highest precipitation fraction throughout the profile for the stratiform threshold. At the same time, the Morrison scheme produces the second lowest cloud fraction for the convective threshold

165 threshold. Given that the total precipitation of the Morrison scheme is not significantly different from the other schemes, this shows that the Morrison scheme produces a much larger area of precipitation with a much lower intensity in general. Another interesting observation is the P3 scheme (purple) which closely follows the observed fraction for both thresholds at upper heights, but starts to deviate at lower heights. This is especially pronounced at the convective threshold (dashed lines)

170 and indicates that the transition from ice to rain seems to be problematic, producing a too high convective precipitation fraction at lower heights of about twice the observed fraction. This also matches the large number of convective cells identified in the P3 scheme (Figure 2). In the following, we focus our analysis on two height levels, 1500 m and 5500 m, indicated with the

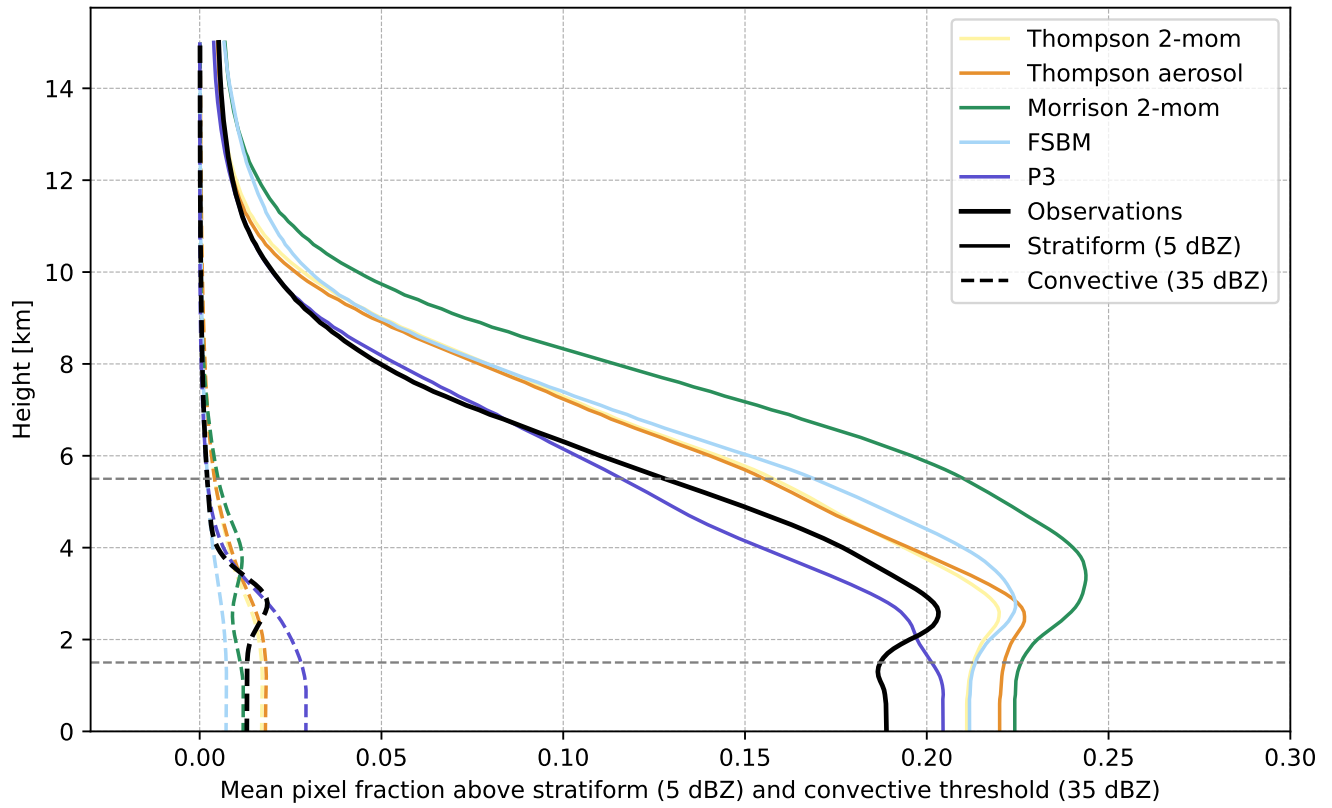


Figure 3. Vertical distribution of the mean pixel fraction larger than a 35 dBZ (dashed) and a 5 dBZ (continuous) threshold.

dashed horizontal lines in Figure 3. This way we cover both, an upper height where we expect to observe differences in ice growth processes, as well as a lower height where we mainly expect precipitating liquid rain drops while at both heights, there is a significant variation in the fraction between the microphysics schemes and to the observations.

3.2 Distribution of convective and stratiform regions

While all schemes produced similar overall amounts of total precipitation, the distribution of precipitation into convective and stratiform regions varies significantly between the microphysics schemes. This is illustrated in Figure 4, which shows the CAF for the 30 days data set for each of the 5 microphysics schemes and the radar observations.

The CAF is visualized using a boxplot, which shows the median (center line), the interquartile range (coloured box), and the range of the data without outliers (whiskers). Outliers are defined as data points that are more than 1.5 times the interquartile range away from the upper or lower quartile. The interquartile range is the difference between the 75th and 25th percentile of the data. Shown are two heights, 1500 m and 5500 m, to illustrate the vertical distribution of the CAF.

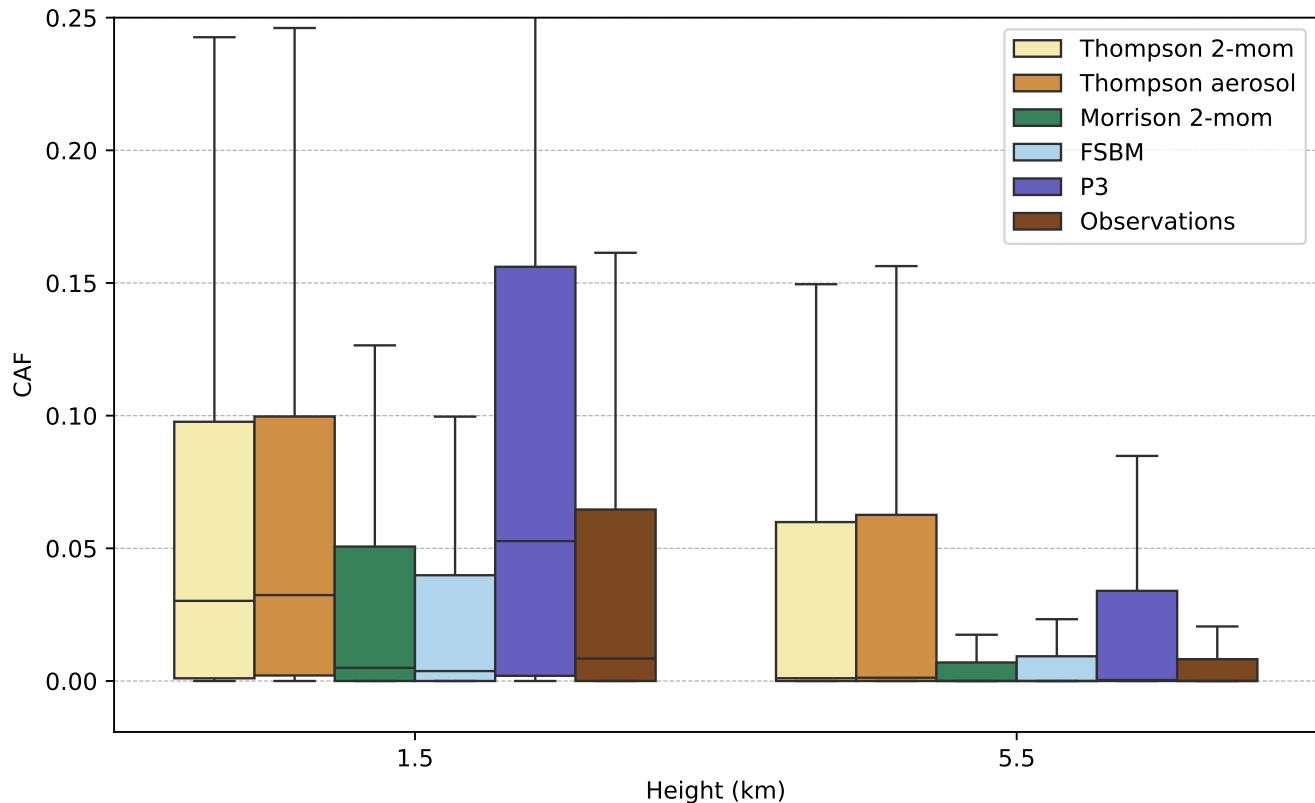


Figure 4. Convective area fraction (CAF) for the 30 days data set for each of the 5 microphysics schemes and the radar observations at two heights, 1500 m and 5500 m.

Generally, the observations have small median CAFs of around 0.8 % at 1500 m. Morrison (0.5 %) and FSBM (0.4 %) are close and only slightly too low. The other 3 schemes have CAFs that are much too high, with the P3 scheme having the highest CAFs (5 %) followed by the Thompson schemes (3 %).

At 5500 m, the CAF is generally lower than at 1500 m, since convective cores not always reach up that high. The overall picture is similar, with a group of schemes (Thompson 2-mom, Thompson aerosol-aware and P3) producing too high CAFs and another group (Morrison and FSBM) producing close and only slightly too low CAFs. Interestingly, while P3 was highest at 1500 m, the Thompson schemes were highest at 5500 m. This is due to the fact that the Thompson schemes produce more frequently high reflectivities at 5.5 km altitude which is attributed to graupel particles (not shown). The distribution into convective and stratiform regions was also evaluated by previous studies. Shrestha et al. (2022) for example found that the CAF was underestimated by their model for 3 convective storm cases. Here, we demonstrate that the choice of microphysics schemes has a significant impact on the distribution of precipitation into convective and stratiform regions. Some schemes,

195 especially the P3 and Thompson schemes produce CAFs that are unrealistically high compared to the observations over the 30 days dataset.

These discrepancies are due to the fact that convective reflectivities of >35 dBZ are produced too frequently in the simulations, especially below the melting layer height within rain. The simulation of these high reflectivities is driven by the underlying rain drop size distribution, which in turn is strongly influenced by microphysical processes. For example, processes
200 such as evaporation, aggregation, or drop break-up can significantly alter the size distribution of hydrometeors and, consequently, the reflectivity. Further insight about the microphysical processes can be gained from the polarimetric radar signals, which are analyzed in the following section.

3.3 Polarimetric radar statistics

Next we will thus have a deeper look into the distributions of reflectivity (Z_{hh}) and, especially, differential reflectivity (Z_{dr})
205 at the two height levels. The figures corresponding to the convective and stratiform regions are shown in Figures 5 and 6, respectively. Shown are histograms of simulated and observed reflectivity (left) and differential reflectivity (right) values for the radar observations and the 5 microphysics schemes. The analysis at 5500 m altitude is shown in the upper row, while the analysis at 1500 m altitude is shown in the lower row. The values are binned into 50 bins (0.6 dBZ and 0.1 dB bins for reflectivity and differential reflectivity, respectively). Below, we provide a detailed analysis of the histograms for stratiform and
210 convective regions separately.

3.3.1 Convective region at 1500 m

The histograms for the convective region at 1500 m altitude are shown in the lower part of the picture. A comparison of the radar observations (black line) and the simulations (colored lines) reveals a significant overestimation of convective (larger than 35 dBZ) reflectivities in the reflectivity distribution (Figure 5 c)) for Thompson 2-moment, Thompson aerosol-aware, and
215 especially the P3 scheme. The FSBM scheme on the other hand has a low bias for most of the reflectivity range, and never produced reflectivities larger than 49 dBZ.

To investigate the reason for the frequency in the reflectivity histogram, the factors that influence radar reflectivity need to be examined. In general, radar reflectivity is influenced by several factors, including the size and number of particles in the beam volume, as well as the density and phase of these particles. At an altitude of 1500 m, the predominant hydrometeor type
220 is rain during the analyzed warm season cases. Consequently, any observed reflectivity bias may be attributable to the size and number of rain drops. It is not possible to distinguish between these potential sources of bias solely based on reflectivity. However, the differential reflectivity (Z_{dr}) can provide further context. Z_{dr} is defined as the ratio of the reflectivity received at the horizontal polarization (Z_{hh}) to the reflectivity received at the vertical polarization (Z_{vv}):

$$Z_{dr} = 10 \log_{10} \left(\frac{Z_{hh}}{Z_{vv}} \right) \quad (1)$$

225 In principle, Z_{dr} is a measure of the aspect ratio of particles in the direction of the radar beam. For rain drops, the aspect ratio is related to their size (Ryzhkov et al., 2011). The resulting Z_{dr} depends on the viewing angle of the radar. From below, rain

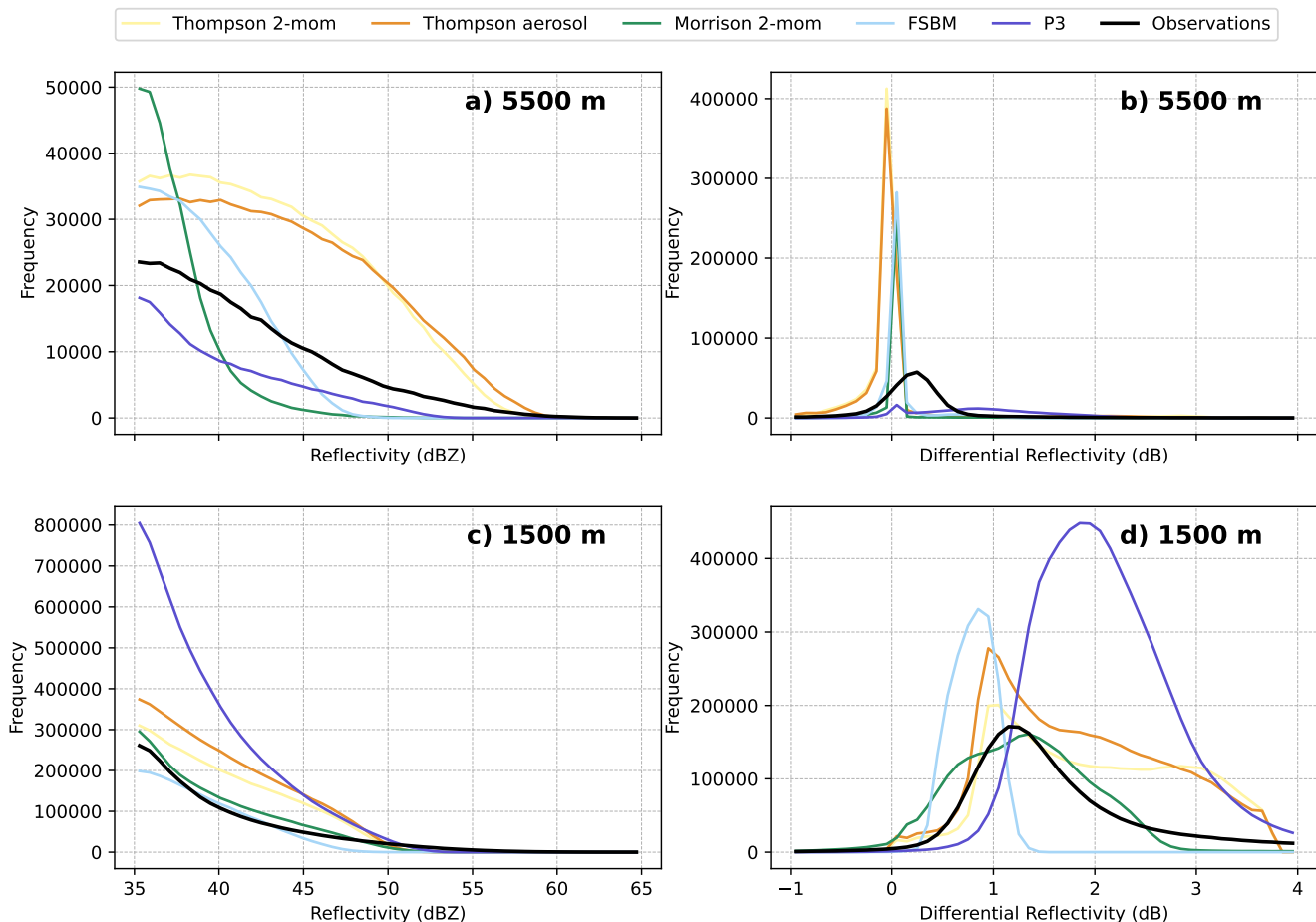


Figure 5. Histograms of simulated and observed reflectivity (left column) and differential reflectivity (right column) for convective precipitation at 5500 m (upper row) and 1500 m (lower row) altitude.

drops appear spherical to the radar, resulting in a Z_{dr} value of 0 dB. At low elevation angles, when the radar looks through the side of the rain particles, rain drops appear more oblate, resulting in Z_{dr} values larger than 0 dB. The magnitude of Z_{dr} is then related to the size of the rain drops, with larger rain drops producing larger Z_{dr} values. Importantly, Z_{dr} is independent of the number of particles in the beam volume. Both, Z_{hh} and Z_{vv} , are proportional to the number of particles in the beam volume, so any change in the number of particles cancels out in the ratio. Consequently, at low elevation angles, Z_{dr} can be used as a measure of rain drop size.

The Z_{dr} histograms for the convective region at 1500 m altitude are shown in Figure 5 d). The observations exhibit a Z_{dr} distribution that has a peak at about 1.2 dB and is skewed to the right, with a tail extending to higher values. In general, the simulations exhibit similar distributions, but with a notable bias towards higher or lower values, depending on the scheme. The FSBM scheme has a strong low bias, generating Z_{dr} values below 1 dB to frequently, while Z_{dr} values above 1.5 dB

are not simulated at all. The Thompson schemes and particularly the P3 scheme demonstrate a high bias in Zdr, simulating high Zdr values of more than 2 dB too frequently. This correlates well with the reflectivity distribution in Figure 5 c). We conclude that the high area fraction of convective precipitation in the P3 and Thompson schemes is due to rain drop sizes that are too large, while the FSBM scheme misses large rain drops and produces too many small rain drops. The Morrison scheme also exhibits low bias in Zdr, though less pronounced. The Morrison scheme exhibits a tendency to simulate Zdr values lower than 1 dB with a higher frequency than observed, yet failed to generate Zdr values greater than 3 dB, despite the existence of such observations. However, this does not translate into a low reflectivity bias. This suggests that large drops were missing in the Morrison scheme relative to the observations, but a large number of droplets compensates for this providing sufficient reflectivity still. Sun et al. (2023) noted a similar lack of large rain drops in the convective part of a squall line in the Morrison scheme. In this study, we demonstrate that this phenomenon is consistent across multiple days and different convective events. However, Sun et al. (2023) reported a low bias of reflectivity in the convective region for the Morrison scheme in their squall line case, which we cannot find based on our simulations. Planche et al. (2019) on the other hand find a high reflectivity bias in the convective part of their squall line case simulation with the Morrison scheme, in line with our findings. This underscores the high variability in the performance of microphysics schemes across different convective events and highlights the strength of a statistical approach.

Relating these findings to the biases in CAF (Figure 4), we see a strong correlation between the high fraction of convective precipitation in the P3 and Thompson schemes and the high bias in the reflectivity and Zdr histograms. Thus, we conclude that the P3 and Thompson schemes produce a convective precipitation area that is too large, which is a result of the frequent occurrence of large rain drops that produce high reflectivities and Zdr values. The FSBM scheme, on the other hand, produces a low fraction of convective precipitation due to the low frequency of large rain drops. This could mean that rain within the P3 and Thompson simulations is generally more intense, but also more localized, while the FSBM scheme produces a more widespread and less intense rain.

3.3.2 Convective region at 5500 m

The reflectivity and differential reflectivity histograms for the convective region at an altitude of 5500 m are shown in Figure 5 a) and b), respectively. These heights provide some context to the preceding analysis conducted at 1500 m altitude, because at mid-latitudes, rain typically originates from melting ice phase particles present at this altitude. Figure 5 a) reveals an overall overestimation of convective (larger 35 dBZ) reflectivities by most schemes. While FSBM and Morrison somewhat compensate for this by a lack of largest values, both Thompson do not. Only P3 shows a deficit of these convective reflectivities in general. The Thompson schemes demonstrate analogous biases at this altitude as observed at 1500 m altitude, e.g., the Thompson schemes exhibit a consistent high bias, generating convective reflectivities too frequently over the entire reflectivity range. This suggests that the issues in the convective core at 1500 m altitude for these schemes stem from the ice phase. For example, in the Thompson schemes, high reflectivities at 1500 m altitude due to rain drop sizes that are too large might be a result of melting ice phase particles that are already too large at 5500 m altitude. Planche et al. (2019) for example show that the rain drop size

270 distribution is very sensitive to the melting process within the Thompson scheme, although their focus was on the melting of
snow.

The Morrison scheme exhibits a notable low bias at 5500 m altitude, resulting in low reflectivities of 39 dBZ and below
being produced too frequently and reflectivities above 39 dBZ being produced too infrequently. This does not translate into
a low reflectivity bias at 1500 m altitude. However, it might explain the low bias in Zdr at 1500 m altitude, as the Morrison
275 scheme produces a low fraction of large rain drops at 1500 m altitude, based on the Zdr histogram (Figure 5 d)).

To provide further context, the Zdr histograms for the convective region at 5500 m altitude are shown in Figure 5 b). At
this altitude, the radar forward operator assumptions regarding the ice shape significantly influence the simulated Zdr. The P3
scheme diverges from the other schemes in the ice phase by employing a single ice class, whereas the other schemes utilize
multiple ice classes. Consequently, the forward operator assumptions differ between the P3 scheme and the other schemes.
280 Further uncertainties arise from the fact that the radar forward operator must assume ice particle shapes as well as particle
orientations, which are not simulated directly by the WRF model. Due to the significant impact of these radar forward operator
assumptions on the differential reflectivity in the ice phase we refrain from further analysis of the Zdr histograms at 5500 m
altitude.

Unfortunately, this means we cannot provide a clear explanation for the bias in reflectivity at 5500 m altitude. However, with
285 the radar forward operator applied in this study (CR-SIM), it is possible to distinguish between the contributions of different
hydrometeor classes to the radar signals. Large reflectivities of more than 30 dBZ at 5500 m altitude are typically associated
with graupel particles. We find that the P3 scheme produce graupel-induced reflectivities of only up to 30 dBZ, Morrison and
FSBM schemes of up to 40 dBZ, while the Thompson schemes frequently yield graupel reflectivities exceeding 50 dBZ (not
shown). We conclude that the bias in reflectivity in the convective core is attributable to graupel particles, where graupel is
290 underrepresented in the P3, Morrison and FSBM scheme, but overrepresented in the Thompson schemes. However, we cannot
distinguish between the effects of number concentration, density assumption or graupel size on the reflectivity bias. Graupel
was also identified as a potential important source of bias in other studies. Sun et al. (2023) for example proposed an underpre-
diction of ice-phase processes associated with graupel and hail, such as riming, as the likely cause for the aforementioned lack
of large rain drops in the Morrison scheme within the context of a squall line case. A different study by Shrestha et al. (2022)
295 revealed that melting of graupel was the dominant rain mechanism in COSMO simulations of 3 convective storms, whereas
the observations indicated a different mechanism.

3.3.3 Stratiform region at 1500 m

Figure 6 illustrates the histograms for the stratiform region. At an altitude of 1500 m, the majority of the schemes exhibit a low
bias in reflectivity (Figure 6 c)). This phenomenon is particularly evident in the Morrison and FSBM schemes, who simulate
300 reflectivity values below 25 dBZ too frequently and reflectivity values above 25 dBZ too infrequently. The Thompson schemes,
produce a peak at the correct reflectivity; however, all schemes produce low reflectivities much too often. This was of course
also evident in Figure 3. Only the P3 scheme distribution is very similar to the observed distribution, however, the frequency
over the entire range is slightly too high.

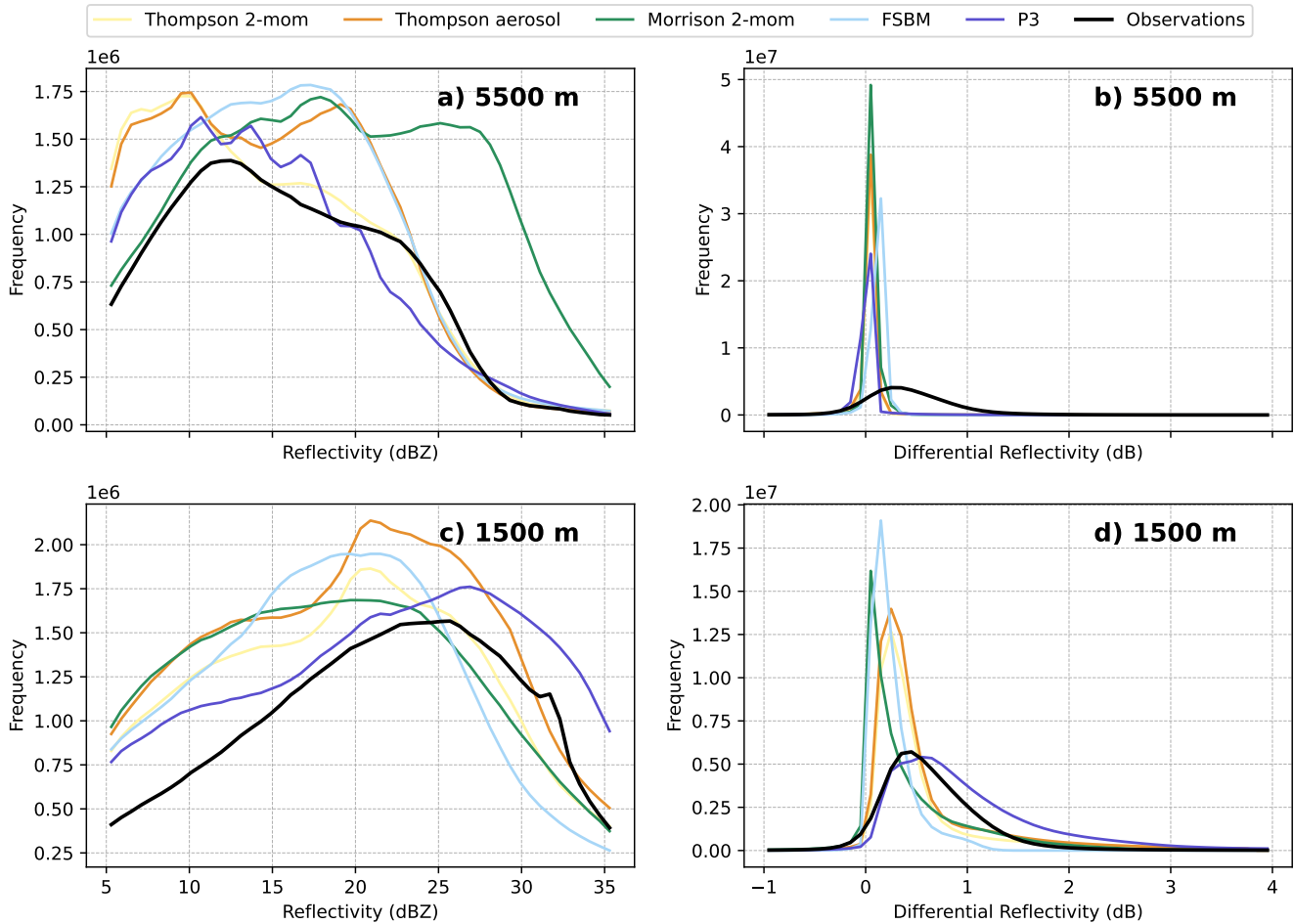


Figure 6. Histograms of simulated and observed reflectivity (left column) and differential reflectivity (right column) for stratiform precipitation at 5500 m (upper row) and 1500 m (lower row) altitude.

Once more, Zdr provides further context (Figure 6 d)). In this case, all schemes except for the P3 scheme demonstrate a pronounced low bias in Zdr, where Zdr below 0.3 – 0.5 dB is simulated too frequently and higher Zdr is simulated too rarely. This suggests that in the stratiform region at 1500 m, all schemes except the P3 scheme simulate large rain drops less often than observed. This pattern is consistent with the convective region at 1500 m altitude for the Morrison and FSBM schemes, where a low bias in Zdr is also observed and is attributed again to a low frequency of large rain drops. While our findings in the convective regions were in line with the findings of Sun et al. (2023), the results in the stratiform region are not. Sun et al. (2023) reported an excessive proportion of large drops for the Morrison scheme for the stratiform regions of their squall line case, whereas we find that over a 30-day dataset, the Morrison scheme exhibits a low bias in Zdr in the stratiform regions. A

possible explanation for this discrepancy could be that the large drops that Sun et al. (2023) find in the stratiform region of their squall line is a case specific phenomenon, which is not a general bias of the Morrison scheme.

315 The Morrison and FSBM distributions exhibited similar characteristics both in the stratiform and convective regions at 1500 m altitude, e.g., largest Zdr values were simulated too rarely. In contrast, the Thompson schemes exhibit a completely different behavior in the stratiform region compared to the convective region at 1500 m altitude. There is a clear overestimation of large Zdr values in the convective region, while in the stratiform region, the Thompson schemes exhibit a low bias in Zdr. This could indicate that large particles might fall out within the convective region too efficiently, resulting in a low amount of large particles transported via the ice phase into the stratiform region.

320 Previous studies reported contradictory results regarding rain size biases in the Thompson schemes. Wu et al. (2021) found that their simulation of a typhoon event with the Thompson 2-mom scheme produced a frequency of large rain drops lower than observed. Putnam et al. (2016) on the other hand found that the Thompson 2-mom scheme produced reflectivity values that are too high and attributed this among other things to a high frequency of large rain drops. Our findings suggest that the Thompson schemes produce a relatively high frequency of large rain drops in the convective region, but a relatively low frequency of large
325 rain drops in the stratiform region, which might explain the contradictory findings in previous literature. The P3 scheme is the only scheme that exhibits a Zdr distribution that closely resembles the observations, with only a marginal shift towards higher Zdr values, suggesting that the frequency of large rain drops is more prevalent in the P3 scheme than what is observed. This is consistent with the findings in the convective region at 1500 m altitude, where the P3 scheme also produced a high frequency of large rain drops. However, this bias is much less pronounced in the stratiform region than in the convective region at 1500 m
330 altitude.

3.3.4 Stratiform region at 5500 m

The reflectivity histograms for the stratiform region at an altitude of 5500 m (Figure 6 a) reveal that all schemes produce too large areas of the shown reflectivity levels, but that they do not exhibit a substantial bias. In general, stratiform reflectivities are simulated more frequently than what was observed over most of the reflectivity range for all schemes except the P3 scheme.
335 This is most pronounced for weaker reflectivities below 25 dBZ for all schemes except for the Morrison scheme. The Morrison scheme exhibits a notable deviation of approximately 5 dB towards higher reflectivity values. This finding has been reported in other studies and appears to be a persistent issue with the Morrison scheme. For instance, Planche et al. (2019) reported overestimations of simulated reflectivity of more than 10 dB between 5 and 10 km in their squall line case, attributing this to either snow density assumptions or an excessive number of large snow particles. Sun et al. (2023) also found that the Morrison
340 scheme produced reflectivity values that were too high in the stratiform region below the -20 °C level for their squall line case and attribute this to high assumed graupel densities or too large graupel fractions. In our simulations, we find that reflectivities larger than 25 dBZ at 5.5 km altitude are primarily associated with snow particles in the Morrison scheme simulations (not shown).

Interestingly, the correlation of biases between 1500 m and 5500 m altitude is significantly lower for the stratiform region
345 than for the convective region. The Morrison scheme even has a pronounced high bias at 5500 m altitude, contrasting with

the low bias at 1500 m. This suggests that the origin of biases at 1500 m may not be exclusively attributed to the ice phase for the Morrison scheme within the stratiform region. The disparities in the drop size distributions within the stratiform region may also be attributable to the representation of warm rain processes, such as collisional breakup, collision-coalescence or evaporation. For instance, Planche et al. (2019) reported an unrealistic strong vertical change of drop size diameter and number concentration in the stratiform region of the Morrison simulation of a squall line, due to the representation of warm rain processes in the Morrison scheme. Sun et al. (2023) further reported an overly strong collision-coalescence process in the Morrison scheme for their squall line case in the stratiform region.

The Zdr histograms for the stratiform region at 5500 m altitude are displayed in Figure 6 b). It is evident that no simulation was capable of reproducing the observed Zdr distribution, as all of them exhibited a pronounced peak at approximately 0 dB and almost no values above 0.1 dB. In contrast, the observed distribution is much broader, encompassing values up to 1 dB more frequently. The discrepancy in the Zdr histograms at 5500 m altitude can be attributed to the inherent assumptions of the radar forward operator concerning the particle shapes. Given the similarity in bias observed among all schemes, we refrain from further interpretation of the Zdr histograms at 5500 m altitude.

3.4 Simulated rain drop size distributions

In the previous section, issues in the simulated reflectivity and differential reflectivity at 1500 m altitude were mainly attributed to the size of rain drops. While we do not have real observations of the size distribution of rain drops, we can analyze the simulated size distributions of rain drops in the model output. This helps to understand if the differences in the simulated reflectivity and differential reflectivity are indeed due to differences in the size distribution of rain drops. Figure 7 shows the total number of simulated rain drop sizes over all 30 days at 1500 m altitude for the 5 microphysics schemes for the convective and stratiform regions. The size distributions of the bulk schemes are binned into 17 bins that correspond to the bins used in the FSBM scheme. In general, the distributions within the convective regions and stratiform regions look similar. The area covered by stratiform precipitation is generally larger than the area covered by convective precipitation (see Figure 3). Therefore, the total number of drops in the stratiform region is also generally larger than in the convective region. However, the difference in the number of drops between the two regions is not constant across all drop sizes. For small drops, the number of drops in the stratiform region is significantly larger than in the convective region, by up to 1-2 orders of magnitude. In contrast, the number of large drops of about 4 mm is similar in both, the convective and stratiform regions. This demonstrates that within the convective region, the drop size distribution is shifted towards larger drops, while in the stratiform region, the distribution is shifted towards smaller drops.

The microphysics schemes also impact the size distribution of rain drops. The Morrison and FSBM schemes produce a higher number of small drops, and the lowest number of large drops. The P3 scheme is the exact opposite, producing the highest number of large drops and the lowest number of small drops, while the Thompson schemes are in between. This is true for both regions, convective as well as stratiform, but the differences are more pronounced in the stratiform region, especially for small drops. This is consistent with the interpretations in the previous sections. The schemes that produced the highest Zdr values (Thompson and P3) also produced the highest number of large drops in the size distribution, while the schemes that

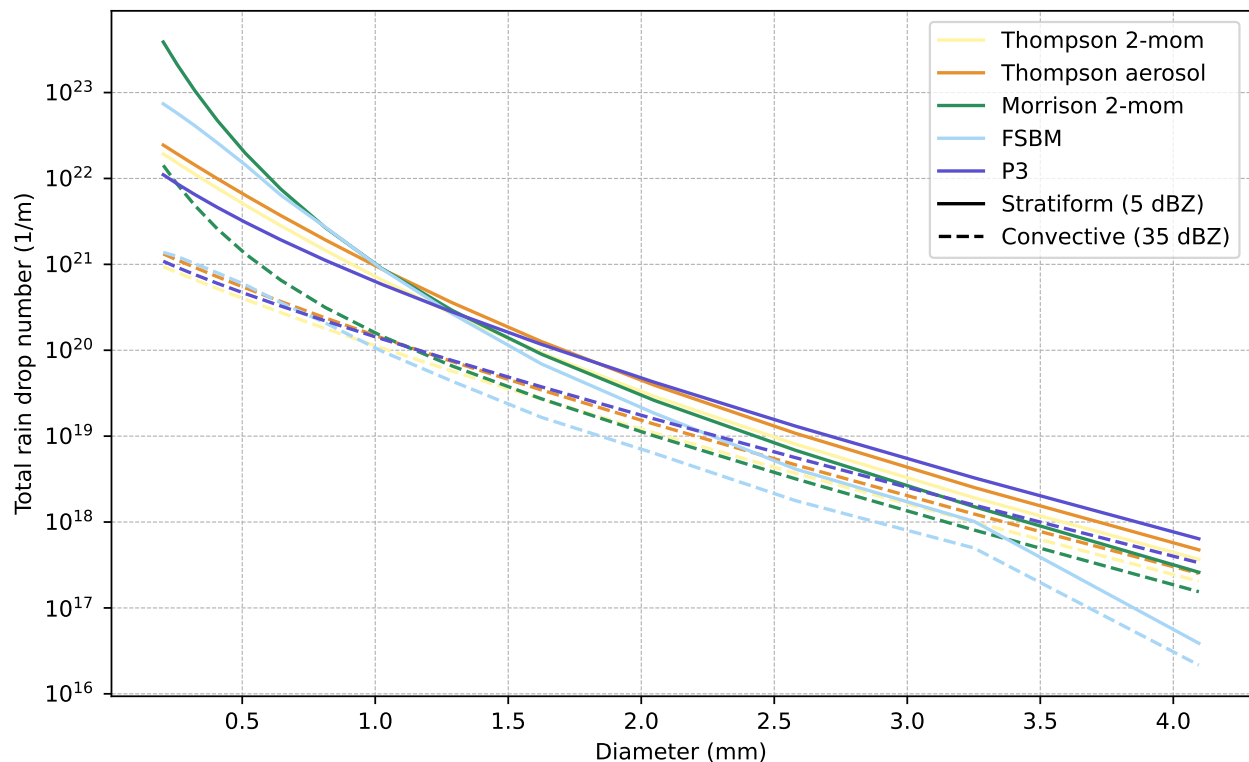


Figure 7. Total simulated number of rain drops for the convective (dashed lines) and stratiform (continuous lines) regions at 1500 m altitude for the 5 microphysics schemes.

380 produced the lowest Z_{dr} values (Morrison and FSBM) also produced the lowest number of large drops in the size distribution. This provides confidence in our conclusions regarding the rain drop size distributions based on the simulated differential radar reflectivity Z_{dr} .

4 Summary and conclusions

The representation of cloud microphysics in numerical weather prediction models are subject to large uncertainties (Morrison et al., 2020; Fan et al., 2017). Particularly challenging is the prediction of convective precipitation events, due to their small scale and rapid error growth of initial uncertainties (Selz and Craig, 2015; Hohenegger and Schär, 2007). In this study, the performance of five different microphysics schemes in the simulation of convective precipitation events is analyzed statistically over a 30-day dataset in 2019 and 2020. The simulations are conducted using the Weather Research and Forecasting (WRF) model, employing a regional limited area model setup over Munich with a horizontal grid spacing of 400 m. 5 different microphysics schemes are utilized: the Thompson 2-moment scheme (Thompson 2-mom; Thompson et al., 2008), the Morrison

385

390

2-moment scheme (Morrison 2-mom; Morrison et al., 2009), the Thompson aerosol-aware scheme (Thompson aerosol-aware; Thompson and Eidhammer, 2014), the fast spectral bin microphysics scheme (FSBM; Shpund et al., 2019), and the Predicted Particle Properties scheme (P3; Morrison and Milbrandt, 2015). The simulations are then compared to polarimetric radar observations from the German Meteorological Service (DWD) radar network, i.e., reflectivity (Z) and differential reflectivity (Zdr).
395 The convective and stratiform regions are separated using an automatic cell-tracking algorithm, *tobac* (Sokolowsky et al., 2024).

The convective area fraction (CAF) is a metric employed to assess the distribution of precipitation into convective and stratiform regions. While the total precipitation amounts are comparable for all schemes, the distribution into convective and stratiform regions varies significantly. The Thompson schemes and P3 exhibit higher CAFs compared to the observed values,
400 while the Morrison schemes and FSBM demonstrate close, though slightly lower CAFs. The discrepancies are mainly a result of differences in the simulated convective area, rather than the stratiform area. This underscores the pivotal role that the choice of microphysics schemes plays in determining the distribution of precipitation into convective and stratiform regions.

To investigate microphysical reasons for the differences in precipitation distributions, an analysis of simulated and observed reflectivity and differential reflectivity histograms is conducted for the convective and stratiform regions at two altitudes,
405 1500 m and 5500 m. Polarimetric radar signals are influenced by the size and number of the particles in the beam volume, as well as the density and phase of the particles. This allows for an interpretation of the biases in reflectivity and differential reflectivity in terms of the microphysical properties of the simulated particles. The following conclusions are derived from the analysis of the radar signal histograms:

410 **Convective Core**

1. At an altitude of 1500 m, FSBM and Morrison simulations frequently lack large rain drops. For the FSBM scheme, this results in a low bias in reflectivity. For the Morrison scheme this is not the case as it is compensated by a high number of small drops.
- 415 2. The Thompson schemes and P3 produce an excessive number of high Zdr values, indicating that large rain drops are too frequent in these schemes. This is consistent with the simulated rain drop size distributions.
3. The Thompson schemes exhibit comparable biases in reflectivity at 1500 m and 5500 m altitude, indicating that these biases originate in the ice phase.

Sun et al. (2023) also reported a lack of large rain drops in the convective part of a squall line in the Morrison scheme, a
420 finding consistent with our results. Here, we demonstrate that this is persistent across multiple days and different convective events. Contradictory results have been reported in other studies concerning the reflectivity biases of the Morrison scheme. Sun et al. (2023), for instance, reported a low bias within the convective region for their squall line case, while Planche et al. (2019) found a high bias in the convective part for their squall line case. This demonstrates the high variability of the performance of

microphysics schemes for different convective events and the strength of a statistical approach.

425

Stratiform region

1. Only the P3 scheme produces a sufficient number of large rain drops, while all other schemes produce too few large rain drops, resulting in a low bias in Zdr and reflectivity.
- 430 2. At an altitude of 5500 m, the Morrison scheme exhibits a significant high bias in reflectivity.
3. There is no significant correlation between distributions at 1500 m and 5500 m altitude, indicating that the origin of the bias is not necessarily in the ice phase.

Other studies also reported that the Morrison scheme yielded reflectivity values that were excessively high in the stratiform region at upper heights. Planche et al. (2019), for instance, reported overestimations of Morrison's simulated reflectivity of more
435 than 10 dB between 5 and 10-km altitude for their squall line case. They attributed this to either snow density assumptions or too many large snow particles. Sun et al. (2023) also found that the Morrison scheme produced reflectivity values that were too high in the stratiform region below the -20 °C level for their squall line case and attributed this to graupel densities or too large graupel fractions.

In general, the Morrison and FSBM schemes exhibit similar biases in Zdr in the stratiform and convective regions at 1500 m
440 altitude, i.e., low Zdr values are simulated too frequently in both regions. The Thompson schemes on the other hand have a bias in Zdr at 1500 m altitude in the convective region, but a low bias in Zdr at 1500 m altitude in the stratiform region. This could indicate that large particles might sediment too efficiently within the convective region, resulting in a low amount of large particles transported via the ice phase into the stratiform region. Previous studies reported contradictory results regarding rain size biases in the Thompson schemes. For example, Wu et al. (2021) reported a frequency of large drops lower than observed,
445 while Putnam et al. (2016) found that the Thompson 2-mom scheme produced reflectivity values that are too high and attributed this to a high frequency of large rain drops. Our findings suggest that the Thompson schemes do not generally exhibit a size bias in rain drops, but that the size bias is dependent on the region. This might explain the contradictory findings in previous literature.

Relating the analysis based on the polarimetric radar signals to the precipitation distribution based on the CAF, we find a
450 strong correlation between the high fraction of convective precipitation in the P3 and Thompson schemes and the high bias in reflectivity and Zdr histograms. We conclude that the high CAF in the P3 and Thompson schemes is due to the high frequency of large rain drops that produce high reflectivities and Zdr values. This translates into an unrealistic high fraction of convective precipitation in the P3 and Thompson schemes. With the same argument, we conclude that the missing large drops in the FSBM and Morrison scheme translate into a low CAF, although this is not as pronounced as in the P3 and Thompson schemes.

455 In general, the majority of the observed biases in reflectivity and Zdr can be attributed to issues in the simulation of particle size distributions. Analysis of the underlying simulated rain drop size distributions (DSDs) supports these interpretations of the

Zdr signals. This is of particular significance, because the representation of particle size distributions is crucial for the accurate simulation of precipitation.

Our study shows that the choice of microphysics schemes has a significant impact on the rain drop size distributions, which in turn affects the distribution of precipitation into intense localized convective and more widespread, weaker stratiform regions. We demonstrate how polarimetric radar observations can be used to identify issues in the underlying particle size distributions of the microphysics schemes. In contrast to previous studies that evaluate microphysics schemes typically based on case studies, our evaluation encompasses a dataset over 30 days of convective precipitation events. This provides a more robust estimate of the performance of microphysics schemes. In light of the increasing accessibility of operational polarimetric radar observations, our study has the potential to provide a framework for the rigorous statistical evaluation of weather models using polarimetric radar measurements, extending this capability to other regions and weather scenarios. However, within the scope of this study, we focused only on one polarimetric radar variable, Zdr. Future research could expand the analysis to include additional polarimetric radar variables, such as specific differential phase (Kdp) or correlation coefficient (RHO_{hv}), which could provide further insights into the microphysical properties of the different microphysics schemes. Kdp, for instance, can provide information on number concentration of rain drops in the radar beam volume, while RHO_{hv} can provide information on the presence of mixed-phase precipitation (Bringi and Chandrasekar, 2001, chapter 7). Kdp and RHO_{hv} in combination with Zdr and Z could also be used for hydrometeor classification (e.g., Dolan et al., 2013).

While polarimetric radar measurements can provide valuable insights into the microphysical properties of hydrometeors, exact conclusions regarding the DSDs could be derived through the use of disdrometer measurements or retrievals, such as Dual-Wavelength retrievals. Li et al. (2023), for instance, evaluated a simulation of an extreme rainfall event over China using disdrometer measurements and underscored the importance of drop size distributions when analyzing extreme rainfall events.

Furthermore, the radar forward operator assumptions on the particle shapes exert a substantial influence on the simulated Zdr for ice phase particles. This hinders the interpretation of the Zdr histograms at 5500 m altitude in this study. The use of a more sophisticated radar forward operator that does not rely on the T-matrix method could provide more insights into the microphysical properties of the different microphysics schemes. Alternatively, the utilization of schemes that provide more detailed information on the particle shapes and orientations could mitigate the uncertainties in the radar forward operator assumptions.

Finally, this study did not consider cell lifetime stage. However, with cell tracking tools such as tobac, it is possible to obtain further insights into the microphysical properties of the different microphysics schemes during various stages of a convective cell's lifetime. This could be a topic for future research.

Code and data availability. The code for the tobac cell tracking algorithm is available on GitHub at <https://github.com/tobac-project/tobac> (last access 7 April 2025; Sokolowsky et al., 2024). The Weather Research and Forecasting model (WRF, version 4.2) is publicly available on GitHub at <https://github.com/wrf-model/WRF> (last access: 7 April 2025; <https://doi.org/10.5065/1dfh-6p97>, Skamarock et al., 2019). The forward operator CR-SIM (version 3.33) is available on the website of Stony Brook University: <https://you.stonybrook.edu/radar/research/radar->

490 simulators/ (last access 7 April 2025; Oue et al., 2020). The original polarimetric radar data from the operational C-band radar in Isen are
available for research from the German Weather Service (DWD) upon request. The dataset that was used in this work for evaluation, e.g.,
simulations and observations after interpolation to a regular grid is available at: <https://opendata.physik.lmu.de/2rvyDFh8t2W4QNX/> . The
original WRF simulations are available at: <https://opendata.physik.lmu.de/gmjQaVOKkmbQCqC/> . Both these datasets will be made available
with a doi on the opendata server of the LMU (<https://opendata.physik.lmu.de/>) after a successful review of this manuscript. The software
495 that was developed to conduct the analysis in this study is available on zenodo at <https://doi.org/10.5281/zenodo.15519860>.

Author contributions. GK developed the methodology presented in this study, conducted the simulations and the analysis, and wrote the
manuscript. TZ supervised the study, discussed the scientific content, and commented on the manuscript.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. This research has been supported by the project IcePolCKa (Investigation of the initiation of convection and the evolu-
500 tion of precipitation using simulations and polarimetric radar observations at C- and Ka-band) funded by the Deutsche Forschungsgemein-
schaft (DFG, German Research Foundation) – 408027579 – as part of the special priority program on the Fusion of Radar Polarimetry and
Atmospheric Modelling (DFG SPP-2115, PROM). AI tools (Github Copilot, deepl.com) were used to assist the code development and for
language improvements. All ideas, methods, and conclusions presented in this manuscript were developed without the use of AI tools.

References

- 505 Allan, D. B., Caswell, T., Keim, N. C., van der Wel, C. M., and Verweij, R. W.: soft-matter/trackpy: Trackpy v0.5.0, <https://doi.org/10.5281/ZENODO.4682814>, 2021.
- Austin, P. M. and Bemis, A. C.: A QUANTITATIVE STUDY OF THE “BRIGHT BAND” IN RADAR PRECIPITATION ECHOES, *Journal of Atmospheric Sciences*, 7, 145 – 151, [https://doi.org/10.1175/1520-0469\(1950\)007<0145:AQSOTB>2.0.CO;2](https://doi.org/10.1175/1520-0469(1950)007<0145:AQSOTB>2.0.CO;2), 1950.
- Barber, P. and Yeh, C.: Scattering of electromagnetic waves by arbitrarily shaped dielectric bodies, *Applied Optics*, 14, 2864, 510 <https://doi.org/10.1364/ao.14.002864>, 1975.
- Billault-Roux, A.-C., Grazioli, J., Delanoë, J., Jorquera, S., Pauwels, N., Viltard, N., Martini, A., Mariage, V., Le Gac, C., Caudoux, C., Aubry, C., Bertrand, F., Schwarzenboeck, A., Jaffeux, L., Coutris, P., Febvre, G., Pichon, J. M., Dezitter, F., Gehring, J., Untersee, A., Calas, C., Figueras i Ventura, J., Vie, B., Peyrat, A., Curat, V., Rebouissoux, S., and Berne, A.: ICE GENESIS: Synergetic Aircraft and Ground-Based Remote Sensing and In Situ Measurements of Snowfall Microphysical Properties, *Bulletin of the American Meteorological Society*, 104, E367–E388, <https://doi.org/10.1175/bams-d-21-0184.1>, 2023.
- 515 Bringi, V. N. and Chandrasekar, V.: *Polarimetric Doppler Weather Radar*, Cambridge University Press, ISBN 9780521623841, <https://doi.org/10.1017/cbo9780511541094>, 2001.
- Brown, B. R., Bell, M. M., and Frambach, A. J.: Validation of simulated hurricane drop size distributions using polarimetric radar, *Geophysical Research Letters*, 43, 910–917, <https://doi.org/10.1002/2015gl067278>, 2016.
- 520 Caine, S., Lane, T. P., May, P. T., Jakob, C., Siems, S. T., Manton, M. J., and Pinto, J.: Statistical Assessment of Tropical Convection-Permitting Model Simulations Using a Cell-Tracking Algorithm, *Monthly Weather Review*, 141, 557–581, <https://doi.org/10.1175/mwr-d-11-00274.1>, 2013.
- Chen, G., Zhao, K., Huang, H., Yang, Z., Lu, Y., and Yang, J.: Evaluating Simulated Raindrop Size Distributions and Ice Microphysical Processes With Polarimetric Radar Observations in a Meiyu Front Event Over Eastern China, *Journal of Geophysical Research: Atmospheres*, 525 126, <https://doi.org/10.1029/2020jd034511>, 2021.
- Chen, X. and Liu, X.: Comparison of the Morrison and WDM6 Microphysics Schemes in the WRF Model for a Convective Precipitation Event in Guangdong, China, Through the Analysis of Polarimetric Radar Data, *Remote Sensing*, 16, 3749, <https://doi.org/10.3390/rs16193749>, 2024.
- Dolan, B., Rutledge, S. A., Lim, S., Chandrasekar, V., and Thurai, M.: A Robust C-Band Hydrometeor Identification Algorithm and Application to a Long-Term Polarimetric Radar Dataset, *Journal of Applied Meteorology and Climatology*, 52, 2162–2186, 530 <https://doi.org/10.1175/jamc-d-12-0275.1>, 2013.
- Doswell III, C. A.: Severe Convective Storms – An Overview, in: *Severe Convective Storms*, chap. 1, pp. 1 – 26, American Meteorological Society, ISBN 9781878220417, 2001.
- Fan, J., Han, B., Varble, A., Morrison, H., North, K., Kollias, P., Chen, B., Dong, X., Giangrande, S. E., Khain, A., Lin, Y., Mansell, E., 535 Milbrandt, J. A., Stenz, R., Thompson, G., and Wang, Y.: Cloud-resolving model intercomparison of an MC3E squall line case: Part I—Convective updrafts, *Journal of Geophysical Research: Atmospheres*, 122, 9351–9378, <https://doi.org/10.1002/2017jd026622>, 2017.
- Flack, D. L. A., Gray, S. L., and Plant, R. S.: A simple ensemble approach for more robust process-based sensitivity analysis of case studies in convection-permitting models, *Quarterly Journal of the Royal Meteorological Society*, 145, 3089–3101, <https://doi.org/10.1002/qj.3606>, 2019.

- 540 Han, B., Fan, J., Varble, A., Morrison, H., Williams, C. R., Chen, B., Dong, X., Giangrande, S. E., Khain, A., Mansell, E., Milbrandt, J. A., Shpund, J., and Thompson, G.: Cloud-Resolving Model Intercomparison of an MC3E Squall Line Case: Part II. Stratiform Precipitation Properties, *Journal of Geophysical Research: Atmospheres*, 124, 1090–1117, <https://doi.org/10.1029/2018jd029596>, 2019.
- Helmert, K., Tracksdorf, P., Steinert, J., Werner, M., Frech, M., Rathmann, N., Hengstebeck, T., Mott, M., Schumann, S., and Mammen, T.: DWDs new radar network and post-processing algorithm chain, in: Proc. Eighth European Conf. on Radar in Meteorology and Hydrology (ERAD 2014), Garmisch-Partenkirchen, Germany, https://www.pa.op.dlr.de/erad2014/programme/ExtendedAbstracts/237_Helmert.pdf, last access: 16 February 2022, 2014.
- Hohenegger, C. and Schär, C.: Predictability and Error Growth Dynamics in Cloud-Resolving Models, *Journal of the Atmospheric Sciences*, 64, 4467–4478, <https://doi.org/10.1175/2007jas2143.1>, 2007.
- Houze, R. A.: Cloud Dynamics, vol. 53 of *International Geophysics*, Academic Press, ISBN 9780123568816, 1994.
- 550 Kumjian, M. R.: The impact of precipitation physical processes on the polarimetric radar variables, The University of Oklahoma, 2012.
- Köcher, G., Zinner, T., Knote, C., Tetoni, E., Ewald, F., and Hagen, M.: Evaluation of convective cloud microphysics in numerical weather prediction models with dual-wavelength polarimetric radar observations: methods and examples, *Atmospheric Measurement Techniques*, 15, 1033–1054, <https://doi.org/10.5194/amt-15-1033-2022>, 2022.
- Köcher, G., Zinner, T., and Knote, C.: Influence of cloud microphysics schemes on weather model predictions of heavy precipitation, *Atmospheric Chemistry and Physics*, 23, 6255–6269, <https://doi.org/10.5194/acp-23-6255-2023>, 2023.
- 555 Lamb, D.: CLOUD MICROPHYSICS, pp. 459–467, Elsevier, ISBN 9780122270901, <https://doi.org/10.1016/b0-12-227090-8/00111-1>, 2003.
- Li, H., Huang, Y., Luo, Y., Xiao, H., Xue, M., Liu, X., and Feng, L.: Does “right” simulated extreme rainfall result from the “right” representation of rain microphysics?, *Quarterly Journal of the Royal Meteorological Society*, 149, 3220–3249, <https://doi.org/10.1002/qj.4553>, 560 2023.
- Morrison, H. and Milbrandt, J. A.: Parameterization of Cloud Microphysics Based on the Prediction of Bulk Ice Particle Properties. Part I: Scheme Description and Idealized Tests, *Journal of the Atmospheric Sciences*, 72, 287–311, <https://doi.org/10.1175/jas-d-14-0065.1>, 2015.
- Morrison, H., Thompson, G., and Tatarskii, V.: Impact of Cloud Microphysics on the Development of Trailing Stratiform Precipitation in a Simulated Squall Line: Comparison of One- and Two-Moment Schemes, *Monthly Weather Review*, 137, 991–1007, <https://doi.org/10.1175/2008mwr2556.1>, 2009.
- 565 Morrison, H., van Lier-Walqui, M., Fridlind, A. M., Grabowski, W. W., Harrington, J. Y., Hoose, C., Korolev, A., Kumjian, M. R., Milbrandt, J. A., Pawlowska, H., Posselt, D. J., Prat, O. P., Reimel, K. J., Shima, S.-I., van Dierenhoven, B., and Xue, L.: Confronting the Challenge of Modeling Cloud and Precipitation Microphysics, *Journal of Advances in Modeling Earth Systems*, 12, <https://doi.org/10.1029/2019ms001689>, 2020.
- 570 National Centers For Environmental Prediction/National Weather Service/NOAA/U.S. Department Of Commerce: NCEP GFS 0.25 Degree Global Forecast Grids Historical Archive, <https://doi.org/10.5065/D65D8PWK>, 2015.
- Oue, M., Tatarevic, A., Kollias, P., Wang, D., Yu, K., and Vogelmann, A. M.: The Cloud-resolving model Radar SIMulator (CR-SIM) Version 3.3: description and applications of a virtual observatory, *Geoscientific Model Development*, 13, <https://doi.org/10.5194/gmd-13-1975-2020>, (code available at: <https://you.stonybrook.edu/radar/research/radar-simulators/>, last access: 21 September 2021), 2020.

- Planche, C., Tridon, F., Banson, S., Thompson, G., Monier, M., Battaglia, A., and Wobrock, W.: On the Realism of the Rain Microphysics Representation of a Squall Line in the WRF Model. Part II: Sensitivity Studies on the Rain Drop Size Distributions, *Monthly Weather Review*, 147, 2811–2825, <https://doi.org/10.1175/mwr-d-18-0019.1>, 2019.
- Powell, S. W., Houze, R. A., and Brodzik, S. R.: Rainfall-type categorization of radar echoes using polar coordinate reflectivity data, *Journal of Atmospheric and Oceanic Technology*, 33, 523–538, <https://doi.org/https://doi.org/10.1175/JTECH-D-15-0135.1>, 2016.
- Putnam, B. J., Xue, M., Jung, Y., Zhang, G., and Kong, F.: Simulation of Polarimetric Radar Variables from 2013 CAPS Spring Experiment Storm-Scale Ensemble Forecasts and Evaluation of Microphysics Schemes, *Monthly Weather Review*, 145, 49–73, <https://doi.org/10.1175/mwr-d-15-0415.1>, 2016.
- Qian, Q., Lin, Y., Luo, Y., Zhao, X., Zhao, Z., Luo, Y., and Liu, X.: Sensitivity of a Simulated Squall Line During Southern China Monsoon Rainfall Experiment to Parameterization of Microphysics, *Journal of Geophysical Research: Atmospheres*, 123, 4197–4220, <https://doi.org/10.1002/2017jd027734>, 2018.
- Ryzhkov, A., Pinsky, M., Pokrovsky, A., and Khain, A.: Polarimetric Radar Observation Operator for a Cloud Model with Spectral Microphysics, *Journal of Applied Meteorology and Climatology*, 50, 873–894, <https://doi.org/10.1175/2010jamc2363.1>, 2011.
- Ryzhkov, A. V., Snyder, J., Carlin, J. T., Khain, A., and Pinsky, M.: What Polarimetric Weather Radars Offer to Cloud Modelers: Forward Radar Operators and Microphysical/Thermodynamic Retrievals, *Atmosphere*, 11, 362, <https://doi.org/10.3390/atmos11040362>, 2020.
- Seidel, J. S., Kiselev, A. A., Keinert, A., Stratmann, F., Leisner, T., and Hartmann, S.: Secondary ice production – no evidence of efficient rime-splintering mechanism, *Atmospheric Chemistry and Physics*, 24, 5247–5263, <https://doi.org/10.5194/acp-24-5247-2024>, 2024.
- Seifert, A. and Beheng, K. D.: A two-moment cloud microphysics parameterization for mixed-phase clouds. Part 1: Model description, *Meteorology and Atmospheric Physics*, 92, 45–66, <https://doi.org/10.1007/s00703-005-0112-4>, 2005.
- Selz, T. and Craig, G. C.: Upscale Error Growth in a High-Resolution Simulation of a Summertime Weather Event over Europe*, *Monthly Weather Review*, 143, 813–827, <https://doi.org/10.1175/mwr-d-14-00140.1>, 2015.
- Shpund, J., Khain, A., Lynn, B., Fan, J., Han, B., Ryzhkov, A., Snyder, J., Dudhia, J., and Gill, D.: Simulating a Mesoscale Convective System Using WRF With a New Spectral Bin Microphysics: 1: Hail vs Graupel, *Journal of Geophysical Research: Atmospheres*, 124, 14 072–14 101, <https://doi.org/10.1029/2019jd030576>, 2019.
- Shrestha, P., Trömel, S., Evaristo, R., and Simmer, C.: Evaluation of modelled summertime convective storms using polarimetric radar observations, *Atmospheric Chemistry and Physics*, 22, 7593–7618, <https://doi.org/10.5194/acp-22-7593-2022>, 2022.
- Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Liu, Z., Berner, J., Wang, W., Powers, J. G., Duda, M. G., Barker, D. M., and Huang, X.-Y.: A Description of the Advanced Research WRF Model Version 4, Tech. rep., <https://doi.org/10.5065/1DFH-6P97>, (code available on GitHub at: <https://github.com/wrf-model/WRF>, last access: 20 June 2020), 2019.
- Sokolowsky, G. A., Freeman, S. W., Jones, W. K., Kukulies, J., Senf, F., Marinescu, P. J., Heikenfeld, M., Brunner, K. N., Bruning, E. C., Collis, S. M., Jackson, R. C., Leung, G. R., Pfeifer, N., Raut, B. A., Saleeby, S. M., Stier, P., and van den Heever, S. C.: tobacv1.5: introducing fast 3D tracking, splits and mergers, and other enhancements for identifying and analysing meteorological phenomena, *Geoscientific Model Development*, 17, 5309–5330, <https://doi.org/10.5194/gmd-17-5309-2024>, 2024.
- Stein, T. H. M., Hogan, R. J., Clark, P. A., Halliwell, C. E., Hanley, K. E., Lean, H. W., Nicol, J. C., and Plant, R. S.: The DYMECS Project: A Statistical Approach for the Evaluation of Convective Storms in High-Resolution NWP Models, *Bulletin of the American Meteorological Society*, 96, 939–951, <https://doi.org/10.1175/bams-d-13-00279.1>, 2015.
- Sun, Y., Zhou, Z., Gao, Q., Li, H., and Wang, M.: Evaluating Simulated Microphysics of Stratiform and Convective Precipitation in a Squall Line Event Using Polarimetric Radar Observations, *Remote Sensing*, 15, 1507, <https://doi.org/10.3390/rs15061507>, 2023.

- 615 Tana, G., Ri, X., Shi, C., Ma, R., Letu, H., Xu, J., and Shi, J.: Retrieval of cloud microphysical properties from Himawari-8/AHI infrared channels and its application in surface shortwave downward radiation estimation in the sun glint region, *Remote Sensing of Environment*, 290, 113 548, <https://doi.org/10.1016/j.rse.2023.113548>, 2023.
- Thompson, G. and Eidhammer, T.: A Study of Aerosol Impacts on Clouds and Precipitation Development in a Large Winter Cyclone, *Journal of the Atmospheric Sciences*, 71, 3636–3658, <https://doi.org/10.1175/jas-d-13-0305.1>, 2014.
- 620 Thompson, G., Field, P. R., Rasmussen, R. M., and Hall, W. D.: Explicit Forecasts of Winter Precipitation Using an Improved Bulk Microphysics Scheme. Part II: Implementation of a New Snow Parameterization, *Monthly Weather Review*, 136, 5095–5115, <https://doi.org/10.1175/2008mwr2387.1>, 2008.
- Trömel, S., Simmer, C., Blahak, U., Blanke, A., Doktorowski, S., Ewald, F., Frech, M., Gergely, M., Hagen, M., Janjic, T., Kalesse-Los, H., Kneifel, S., Knote, C., Mendrok, J., Moser, M., Köcher, G., Mühlbauer, K., Myagkov, A., Pejcic, V., Seifert, P., Shrestha, P., Teisseire, A., von Terzi, L., Tetoni, E., Vogl, T., Voigt, C., Zeng, Y., Zinner, T., and Quaas, J.: Overview: Fusion of radar polarimetry and numerical atmospheric modelling towards an improved understanding of cloud and precipitation processes, *Atmospheric Chemistry and Physics*, 21, 17 291–17 314, <https://doi.org/10.5194/acp-21-17291-2021>, 2021.
- 625 Wu, D., Zhang, F., Chen, X., Ryzhkov, A., Zhao, K., Kumjian, M. R., Chen, X., and Chan, P.-W.: Evaluation of Microphysics Schemes in Tropical Cyclones Using Polarimetric Radar Observations: Convective Precipitation in an Outer Rainband, *Monthly Weather Review*, 149, 1055–1068, <https://doi.org/10.1175/mwr-d-19-0378.1>, 2021.
- 630 Xue, L., Fan, J., Lebo, Z. J., Wu, W., Morrison, H., Grabowski, W. W., Chu, X., Geresdi, I., North, K., Stenz, R., Gao, Y., Lou, X., Bansemer, A., Heymsfield, A. J., McFarquhar, G. M., and Rasmussen, R. M.: Idealized Simulations of a Squall Line from the MC3E Field Campaign Applying Three Bin Microphysics Schemes: Dynamic and Thermodynamic Structure, *Monthly Weather Review*, 145, 4789–4812, <https://doi.org/10.1175/mwr-d-16-0385.1>, 2017.
- 635 You, C.-R., Chung, K.-S., and Tsai, C.-C.: Evaluating the Performance of a Convection-Permitting Model by Using Dual-Polarimetric Radar Parameters: Case Study of SoWMEX IOP8, *Remote Sensing*, 12, 3004, <https://doi.org/10.3390/rs12183004>, 2020.