Evaluation of convective cloud microphysics in numerical weather prediction model with dual-wavelength polarimetric radar observations: methods and examples

Gregor Köcher[1], Tobias Zinner[1], Christoph Knote[1,3], Eleni Tetoni[2], Florian Ewald[2], and Martin Hagen[2]

[1]Meteorologisches Institut, Ludwig-Maximilians-Universität, Munich, Germany
[2]Deutsches Zentrum für Luft- und Raumfahrt, Institut für Physik der Atmosphäre, Oberpfaffenhofen, Germany
[3]Medizinische Fakultät, Universität Augsburg, Augsburg, Germany

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

Abstract. The representation of cloud microphysical processes contributes substantially to the uncertainty of numerical weather simulations. In part, this is owed to some fundamental knowledge gaps in the underlying processes due to the difficulty to observe them directly. On the path to close these gaps we present a setup for the systematic characterization of differences between numerical weather model and radar observations for convective weather situations. Radar observations are introduced which

5   provide targeted dual-wavelength and polarimetric measurements of convective clouds with the potential to provide more detailed information about hydrometeor shapes and sizes. A convection permitting regional weather model setup is established using 5 different microphysics schemes (double-moment, spectral bin (FSBM), and particle property prediction (P3)). Observations are compared to hindcasts which are created with a polarimetric radar forward simulator for all measurement days. A cell-tracking algorithm applied to radar and model data facilitates comparison on a cell object basis. Statistical comparisons of

10   radar observations and numerical weather model runs are presented on a dataset of 30 convection days. In general, simulations show too few weak and small-scale convective cells. Contoured frequency by altitude distributions of radar signatures reveal deviations between the schemes and observations in ice and liquid phase. Apart from the P3 scheme, simulated reflectivities in the ice phase are too high. Dual-wavelength signatures demonstrate issues of most schemes to correctly represent ice particle size distributions, producing overly large graupel particles. Comparison of polarimetric radar signatures reveal issues of all

15   schemes except the FSBM to correctly represent rain particle size distributions.

1   Introduction

In numerical weather models clouds play an important role by strongly affecting, e.g., the radiation budget or the precipitation formation. Cloud processes are generally divided into two scales: The macrophysics and the microphysics. We refer to "cloud macrophysics" for processes on km scale, namely cloud geometry or cloud coverage, while we refer to "cloud microphysics"

20   for all processes on mm scale or smaller. On coarse-grid weather models, both, macro- and microphysics are unresolved and must be parameterized. Increasing computational power allows numerical weather models to use finer resolutions, which in turn allow to simulate more and more small-scale processes explicitly. Meanwhile some operational weather models partially

resolve convective updrafts (e.g., Pinto et al., 2015; Baldauf et al., 2011; Seity et al., 2011; Lean et al., 2008, and many more). This progress effectively removes problems arising from cloud macrophysical parameterizations, as they can eventually be

25 solved explicitly. However, this is not the case for microphysical processes due to the huge number of hydrometeors present in a cloud. Microphysical processes occur on scales of μm to mm and are not expected to ever be resolved. As such, the parameterization of microphysics in numerical weather simulations is of increasing relative importance the more the model resolution increases.

Although it is well known that cloud microphysics introduce substantial uncertainty to numerical weather simulations (Li

30 et al., 2009; White et al., 2017; Khain et al., 2015; Xue et al., 2017; Morrison et al., 2020, and many more), the extent of this uncertainty and its underlying reasons remain less clear. Microphysical cloud processes are very complex small scale processes, due to the large variety of shapes, sizes and phases of hydrometeors involved. It is a challenge to represent this complexity correctly in a model since it cannot be resolved explicitly. Instead, the effect of the microphysical processes must be parameterized. This has the potential to introduce uncertainties, as important processes could be misrepresented or missed

35 out completely. In numerical weather models different microphysical schemes of varying complexity exist to parameterize the microphysical processes. Traditionally microphysics schemes are categorized into so-called bulk and bin schemes. Bulk schemes assume a pre-defined shape of the particle size distribution of several hydrometeor classes and predict bulk variables, such as the mass mixing ratio for each of the hydrometeor classes. Depending on the predicted number of variables, the scheme is categorized as a 1-moment (e.g., Kessler, 1969), 2-moment (e.g., Morrison et al., 2009) or even 3-moment scheme

40 (e.g., Milbrandt and Yau, 2005). Bin schemes (e.g., Khain et al., 2004) on the other hand do not assume a pre-defined shape of the particle size distribution but instead use a number of size bins and predict the variables for each of the bins independently. In recent years some alternative schemes have been developed: The Predicted Particle Property (P3) scheme (Morrison and Milbrandt, 2015) deviates from partitioning ice particles into categories of hydrometeor classes with corresponding properties but instead predicts the properties of ice particles, such as the riming mass mixing ratio. Lagrangian cloud models (LCM;

45 e.g., Shima et al., 2009) calculate cloud microphysics based on individual particles (super droplet) that represent a family of particles with the same properties, but this type of scheme will not be covered in the present study.

Several studies have analyzed the performance of cloud microphysics schemes by comparing multiple schemes against each other and against observations (Morrison and Pinto, 2006; Gallus Jr and Pfeifer, 2008; Rajeevan et al., 2010; Jankov et al., 2011; Varble et al., 2014; Fan et al., 2015; Li et al., 2015; Fan et al., 2017; Xue et al., 2017; Han et al., 2019, and

50 many more). However, all of them are limited to case studies. Given the large variability between convective cases, a large number of individual cases is necessary to test if one scheme consistently outperform others in reproducing observations (Flack et al., 2019; Stanford et al., 2019). Few studies have evaluated microphysics schemes on such a statistical basis. Johnson et al. (2015) used a statistical emulation approach to study the uncertainty produced by several model input parameters but focused on a single idealized convective cloud simulation. Stein et al. (2015) evaluated simulated convective storms over 40

55 non-consecutive days at varying grid resolutions but with only one microphysics scheme. Caine et al. (2013) describe an object-based approach to statistically compare convective cells of a convection-permitting model with radar observations, but they use only two simple microphysics schemes and their statistics are limited to 4.5 days. By comparing two microphysics

Atmospheric
Measurement
Techniques
Discussions

Open Access

EGU

schemes for different convective events, White et al. (2017) found that the response to cloud droplet number concentrations differs not only between the schemes but also significantly between different convective cases. All of this emphasizes the need

60 for an evaluation of several microphysics schemes over a larger data set on a statistical basis. In an extensive recent overview paper on the challenges in modeling cloud microphysics, Morrison et al. (2020) argue that a rigorous uncertainty quantification on a statistical basis could also help to pinpoint the underlying microphysical processes that cause these uncertainties.

Multiple studies attribute weather simulation errors to poorly constrained cloud microphysics, especially for ice or mixed-phase clouds (e.g., Varble et al., 2014; Stanford et al., 2017). The uncertainty resulting from microphysical cloud processes is

65 in part a result of some fundamental knowledge gaps: It is not well known which processes are poorly represented in numerical models (Morrison et al., 2020). This owes to the difficulty to observe these processes directly. To better constrain the parameters, novel observations are needed to provide corresponding information. These observations must provide information about the key microphysical fingerprints, such as particle properties, their location or ideally conversion rates between hydrometeor classes. Polarimetric radars allow to retrieve hydrometeor classes and shapes and are hence suitable to provide observations

70 of cloud microphysical processes. Within the framework of IcePolCKa (**I**nvestigation of the initiation of **c**onvection and the **e**volution of **p**recipitation using simulati**o**ns and po**l**arimetric radar observations at **C**- and **Ka**-band), a sub-project of the DFG Priority Programme 2115 PROM (Polarimetric Radar Observations meet Atmospheric Modelling - Fusion of Radar Polarimetry and Numerical Atmospheric Modelling Towards an Improved Understanding of Cloud and Precipitation Processes; Trömel et al. (2021)), we exploit the synergy of two polarimetric radars at C- and Ka-band to provide a observational basis for com-

75 parison to numerical weather simulations. We evaluate multiple microphysics of different complexity to answer the question: How much complexity is necessary to reproduce polarimetric radar observations?

The goal of this study is to tackle two different aspects:

1. Provide novel observations of cloud microphysics based on dual-wavelength and polarimetric radar measurements using

80 a combination of operational and research-grade radars

2. Evaluate multiple state-of-the-art cloud microphysics schemes for current generation numerical weather prediction models in a common model framework against observations with a sound statistical basis

Eventually, the evaluation should help to identify microphysical processes with obvious differences between radar measurements and weather simulations. However, it is difficult to extract the influence of the cloud microphysics schemes because of

85 feedbacks between dynamics and microphysics. There are methods that focus on untangling the microphysical impacts from other impacts, e.g., the "piggy backing" method (e.g., Grabowski, 2014). However, operational weather forecast simulations as a whole will always include the feedbacks between microphysics and dynamics as well. Therefore, we decided to use a framework that is applicable to operational weather forecasts and run it over a large number of cases for a statistical comparison, but in this framework we will not be able to perfectly separate the microphysical impacts from possible feedbacks.

90 We present a setup for the systematic characterization of differences between model simulations with different microphysics schemes and polarimetric radar observations for convective weather situations. This includes the application of a radar forward

simulator to the model output and of an automated cell-tracking algorithm to the observations and simulations alike. This allows to objectively compare convective cell characteristics in simulation and observation. We apply this framework to a dataset consisting of 30 days of radar observation and simulations with 5 microphysics schemes of varying complexity.

95 The potential of the generated data set is demonstrated by showing differences in reflectivity between model and observations in convective clouds to identify issues of microphysics schemes to correctly simulate ice and liquid particle size distributions.

The paper is organized as follows. The methods are described in Sect. 2, which includes our radar data (2.1), the simulation setup (2.2), a description of the microphysical schemes (2.3), the radar forward operator used to bring the model output into radar space (2.4), the cell-tracking algorithm (2.5), and the grid matching of the different radars and the model grid

100 (2.6). In Sect. 3 the microphysics schemes are evaluated by comparing statistics of cloud geometry and frequency (3.1) as well as analyzing frequency distributions of reflectivity (3.2), polarimetric variables (3.3), and dual-wavelength ratio (3.4) in simulations and observations. In Sect. 4 the results are discussed.

2   Data and Methodology

In total, we observed and simulated 30 convective days over 2 years in 2019 and 2020. The majority of these days was spring

105 and summer. For all of them convective precipitation was forecasted. A table listing the dates can be found in Appendix A.

2.1   Radar data

The observational data basis is provided by two research radar systems in the area of Munich, Germany, at C- and Ka-band frequency and a complementary third C-band radar operated by the German Weather Service (DWD; Fig. 1). The C-band research radar Poldirad (Schroth et al., 1988), operated by the German Aerospace Center (DLR), is located in Oberpfaffenhofen

110 southwest of Munich. At 23 km distance the research Ka-band radar Mira-35 is operated by the Meteorological Institute Munich (MIM) of the Ludwig-Maximilians-University (LMU) in the center of Munich. The third radar is an operational C-band radar located in Isen at a distance of 40 km to the Mira-35 radar.

All three radars are polarimetric Doppler radars. Poldirad and the Isen radar are fully dual-polarimetric, sending out electromagnetic waves with horizontal and vertical polarization. Both radars receive the co-polar components backscattered by

115 atmospheric targets. Therefore, polarimetric variables such as differential reflectivity (Z_{DR}) or differential phase (K_{DP}) are available. Poldirad additionally receives the cross-polar components and hence measures the linear depolarization ratio (LDR). The Mira-35 radar is a single-polarimetric ground-based cloud radar manufactured by METEK GmbH (Bauer-Pfundstein and Görsdorf, 2007). It only transmits horizontally polarized waves but receives co- and cross-polar components. Thus, it possible to measure LDR in addition to the standard reflectivity.

120 Poldirad and Mira-35 are two research radars without any operational obligations. This allows for synchronized and targeted scan patterns of convective clouds and precipitation on demand. The Isen radar is part of the DWD operational radar network with a fixed observation strategy. For a complete description of the measurement strategy refer to Helmert et al. (2014).

Atmospheric
Measurement
Techniques
Discussions

Open Access

EGU
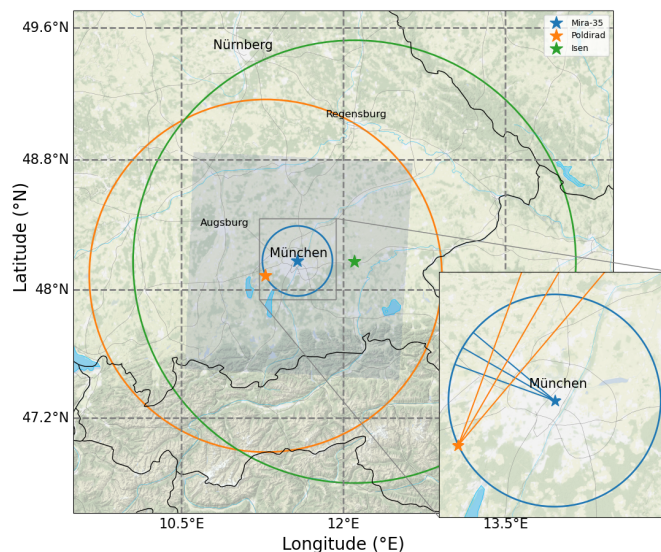
Figure 1. Radar locations and model domain. Filled blue area shows the model domain. Stars show the radar locations and the circles show the radar range around each radar. The straight blue and orange lines visualize RHI scans executed by the Mira-35 and Poldirad radar. Background map tiles by Stamen Design (http://stamen.com), distributed under the Creative Commons Attribution (CC BY 3.0) license. Background map data by OpenStreetMap (http://openstreetmap.org; © OpenStreetMap contributors 2021. Distributed under the Open Data Commons Open Database License (ODbL) v1.0.). Roads, rivers and lakes made with Natural Earth (naturalearthdata.com).

Table 1. Radar characteristics. For the Isen radar the precipitation scan at 1.5° elevation was referenced. For the full configuration of the volume scan, see Helmert et al. (2014).

	POLDIRAD	MIRA-35	ISEN
City	Oberpfaffenhofen	Munich	Isen
Location	48.087 °N, 11.279 °E	48.148 °N, 11.573 °E	48.175 °N, 12.102 °E
Wavelength	5.45 cm	0.85 cm	5.3 cm
Frequency	5.5 GHz	35.2 GHz	5.66 GHz
Beamwidth	1°	0.6°	0.9°
Range	120 km	24 km	150 km

More radar characteristics and configurations can be found in Table 1. This setup allows dual-wavelength and polarimetric measurements of convective clouds and precipitation in the area of Munich.

125   Two measurement strategies have been applied. For spatial coverage data of the operational DWD Isen radar only is utilized in scan strategy A. The Isen radar is running operationally a volume PPI scan every 5 minutes at 11 elevations from 0.5° to 25° and over the whole azimuth circle of 360°. This provides a good spatial coverage at a high temporal resolution. In Figure 1 the green circle depicts the area that is covered by this strategy.

In strategy B, Poldirad and Mira-35 are used for coordinated and targeted scan patterns of the same convective cloud.
130 Strategy B starts with a Poldirad overview scan in plan-position indicator (PPI) mode: The elevation angle is kept constant
and the azimuth angle is varied. After manually choosing a convective cell from this overview PPI, both radars start to execute
three fast scans towards this convective target cloud in the range-height indicator (RHI) scan mode. I.e., the azimuth angle is
kept constant while the elevation angle is varied. The first scan is executed exactly towards the direction that was chosen; one
is directed to 2° azimuth angle to the left; and one is directed to 2° azimuth to the right. This scan mode is referred to as sector
135 range-height indicator (S-RHI). The 9 intersection profiles resulting from these RHIs give an idea about the variation within
the cloud and compensate for potential pointing inaccuracies. In Figure 1 the 6 straight lines (3 orange, 3 blue) visualize these
RHI scans. After each S-RHI scan the azimuth direction is adjusted slightly, according to the projected movement of the cell.
This cell movement is projected using two previous Poldirad overview PPI scans. After a few minutes the S-RHI scans are
stopped (manual) and the procedure starts over with another overview PPI scan. This strategy allows targeted dual-wavelength
140 observations of convective clouds in high vertical resolution over a significant fraction of their life-time.

In total we have collected data of strategy B over 5 convective days during summer 2019. The strategy A comprises a larger
data set. It consists of the same 5 convective days as well as 25 additional convective days during 2019 and 2020.

2.2 Simulation setup

The simulations are performed using the version 4.2 of the Weather Research and Forecasting Model (WRF; Skamarock et al.,
145 2019). Initial and lateral boundary conditions are provided by re-analysis data at 0.25° resolution from the Global Forecast
System (GFS; NCEP, 2015), available every 6 hours and with hourly forecast data in between. Horizontally, the setup includes
a parent Europe domain (3750 km times 3750 km), a two-way nested Germany domain (442 km times 442 km), and a two-way
nested Munich domain (144 km times 144 km). The vertical domain extends from the surface to 5 hPa at 40 vertical levels. The
nesting ratio is 5:1 with the Europe domain at a horizontal resolution of 10 km, the Germany domain at 2 km, and the Munich
150 domain at 400 m. Currently, operational limited area weather models operate at 2 km resolution (e.g., 2.8 km in COSMO-DE
of the German Weather Service; Baldauf et al., 2011). The Munich domain is centered over the Mira-35 instrument (48.15 °N,
11.57 °E). It covers the Mira-35 range (48 km) and an edge region of an additional 48 km around. All analyses are performed on
the innermost Munich domain excluding the edge region, only considering the Mira-35 range (Fig. 1). This area is completely
covered by the Poldirad and Isen radar observations. Each simulation consists of 6 hours spin-up and 24 hours simulation
155 time. The spin-up always starts at 6 pm on the previous day. Thus, the 24 hour forecast exactly covers the day of interest. The
dynamics can freely evolve during the simulation time. Only the parent Europe domain is nudged to the global GFS data. The
inner Germany and Munich domain are not nudged. All days are simulated with 5 different microphysics schemes. Hence,
there are 5 simulations available for each of the convective days and the simulation setups only differ in the choice of the
microphysics scheme.

2.3 Description of microphysics schemes

Five different microphysics schemes are employed. Three 2-moment bulk schemes: One from Thompson et al. (2008) (From here on "Thompson") , the one from (Thompson and Eidhammer, 2014) ("Thompson aerosol-aware") and the one from (Morrison et al., 2009) ("Morrison") as well as the "Fast Spectral Bin Microphysics" (FSBM; Shpund et al., 2019) and the "Predicted Particle Properties" scheme (P3; Morrison and Milbrandt, 2015). The FSBM scheme explicitly resolves the particle size distribution (PSD) with a number of bins, while all other schemes generally represent the PSD by a gamma function

$$N(D) = N_0 D^\mu e^{-\lambda D}, \tag{1}$$

where N_0 is the intercept parameter, D is the particle diameter, μ is the shape parameter, and λ is the slope parameter. The only exception is snow in the Thompson schemes following a bimodal gamma function as described below.

The mass-size relationships are given by a power-law,

$$m = aD^b, \tag{2}$$

where m is the particle mass and D is the particle diameter. The parameter a and b depend on the hydrometeor class and the scheme used and are described below.

a. Thompson

The Thompson bulk scheme predicts integral moments of the PSD for five hydrometeor species: cloud ice, cloud water, rain, snow and graupel. Rain and cloud ice are double-moment species which predict mass mixing ratio (q) and number concentration (N). Snow, graupel and cloud water are single moment, i.e., only the mass mixing ratio is predicted.

The PSDs of rain, cloud ice, graupel and cloud water are represented by gamma distributions (Eq. 1). For rain, graupel and cloud ice $\mu = 0$, i.e., the PSD is an exponential function. Snow and cloud ice have a non-zero μ.

The mass-size relation follows a power-law (Eq. 2). Rain, graupel, cloud ice and cloud water are assumed to be spherical ($b = 3$) with the parameter a depending on the hydrometeor bulk density ρ, with

$$a = \rho \frac{\pi}{6}. \tag{3}$$

The bulk density of rain, graupel, cloud ice and cloud water are constant and size independent.

Snow is treated differently in the Thompson scheme compared to other bulk schemes. Instead of the simple gamma function shown in Eq. 1, a bimodal gamma distribution (sum of an exponential and a gamma function) from Field et al. (2005) that is dependent on temperature is used. Snow is not considered to be spherical. Instead the mass is proportional to D^2 ($b = 2$) to better fit observations. The parameter a of the mass-size relation is constant at $a = 0.069$.

b. Thompson aerosol-aware

190   The Thompson aerosol aware bulk scheme (Thompson and Eidhammer, 2014) is very similar to the older version (Thompson et al., 2008) described in the previous section but includes some changes: While the older version of the Thompson scheme only uses two 2-moment species (rain and cloud ice) and a prescribed number of cloud droplets, the newer version includes activation of aerosols as cloud condensation nuclei (CCN) and ice nuclei (IN). Therefore, it explicitly predicts the droplet number concentration of cloud water and two aerosol variables (CCN and IN).

195

c. Morrison

    The Morrison bulk scheme predicts integral moments of the PSD for five hydrometeor species: cloud ice, cloud water, rain, snow and graupel. All are double moment species. Particle size distributions follow a general gamma distribution (Eq. 1). Rain, cloud ice, snow and graupel have shape parameter $\mu = 0$, again transforming the particle size distributions into an exponential

200 distribution. For cloud water μ is a function of droplet number concentration following Martin et al. (1994). All particles are assumed to be spherical with fixed and size-independent bulk densities.

d. Spectral Bin

    In contrast to the bulk schemes, a spectral bin scheme explicitly resolves the PSD by approximation with a number of

205 independent size bins. This has the advantage that no prior assumption about the shape of the PSD is necessary. However, computational costs are much higher, as all microphysical processes are computed for each bin separately. In this study we use the "Fast Spectral Bin Microphysics" scheme (FSBM; Shpund et al. (2019)) that applies 33 mass-doubling bins, i.e., the mass of the bin k is twice the mass of the bin $k-1$. 5 hydrometeor classes are included: Cloud water, cloud ice, rain, graupel and snow.

210   *e. Predicted Particle Properties (P3)*

    The P3 scheme uses 3 bulk categories: Rain, cloud water and, unlike all the previous schemes, only a single ice category. Instead of predicting mixing ratio and number concentration for multiple ice categories, the P3 scheme predicts properties of this single ice category. Four prognostic ice mixing ratio variables are predicted: total ice mass, rime mass, rime volume and number mixing ratio. Based on these variables more properties are derived, such as rime mass fraction F_r (Ratio of rime mass

215 and ice mass mixing ratio) or rime density ρ_r (Ratio of rime mass and rime volume mixing ratio). All particle size distributions follow a general gamma distribution (Eq. 1). For cloud droplets the shape parameter μ follows observations of Martin et al. (1994). For rain μ follows observations of Cao et al. (2008). For ice μ follows observations of Heymsfield (2003).

    Mass size relationships follow a power law (Eq. 2). The parameter a and b depend on the size of the ice. The scheme distinguishes between small ice, unrimed ice, partially rimed ice and fully rimed ice (graupel/hail). Small ice and graupel are

220 considered spherical ($b = 3$) with parameter a given by Eq. 3, where the ice bulk density ρ equals $917\,\mathrm{kg\,m^{-3}}$ for small ice and varies for graupel/hail. Unrimed ice is considered non-spherical ($b = 1.9$), grown by vapor diffusion or aggregation. The parameter a follows an empirical relationship from Brown and Francis (1995) ($a = 0.0121\,\mathrm{kg\,m^{-b}}$). Partially rimed ice is also considered non-spherical ($b = 1.9$) and the parameter a depends on the rime mass fraction F_r ($a = 0.0121/(1 - F_r)\,\mathrm{kg\,m^{-b}}$),

i.e., a increases with the rime mass fraction. Rain and cloud water are considered spherical with $b = 3$ and a following Eq. (3)

225 and a bulk density ρ of $1000 \ \mathrm{kg\,m^{-3}}$.

2.4 Radar forward operator

To compare the WRF model output against radar observations, version 3.33 of the Cloud Resolving Model Radar Simulator (CR-SIM; Oue et al., 2020) is used. CR-SIM is based on the T-matrix method to compute the scattering characteristics of hydrometeors and is able to simulate polarimetric and Doppler radar variables for several radar frequencies, including C-band

230 and Ka-band that are used in this study. The variables include, among many others, the reflectivity (Z) and specific attenuation (A) at vertical and horizontal polarization, differential reflectivity (Z_{DR}) and specific differential attenuation (ADP). Given that CR-SIM supports both C- and Ka-band frequencies, we are also able to simulate the dual-wavelength ratio (DWR) by performing the forward simulation for the C-band radar as well as the Ka-band radar.

2.5 Cell-Tracking

235 This study focuses on convective clouds and precipitation. To identify and track convective cells in simulations and observations, the open source python package TINT (TINT is not TITAN; Fridlind et al., 2019) is used. TINT is based on the Thunderstorm Identification, Tracking, Analysis and Nowcasting package (TITAN; Dixon and Wiener, 1993). Convective cells are identified using minimum thresholds for reflectivity (32 dBZ) and cell area (8 km^2). A cell motion vector is found by calculating cross-correlation of the reflectivity field in the cell neighborhood of two subsequent time steps and a correction

240 based on prior cell movement. Possible convective cell pairs are compared and matched using an algorithm from TITAN that uses a cost function combining travel distance and volume change of the possible cell pairs. The cell-tracking is applied to simulated and observed reflectivity of the Isen radar only. The simulated and observed reflectivity from Mira-35 and Poldirad is not used for cell-tracking. This way we ensure to have one unique definition to locate convective cells and prevent varying cell definitions depending on the radar that is simulated. More detailed information about TINT can be found in Fridlind

245 et al. (2019) and Dixon and Wiener (1993). TINT does not deal with splits (one cell splits into multiple cells) or mergers (multiple cells merge into one cell), but it was specifically designed for tracking of convective cells over large datasets and is straightforward to apply to our data (Fig. 2).

2.6 Grid Matching and attenuation correction

Radar data and model output are available on different grids. To allow for a comparison these grids must be matched first. In a

250 first step, the model data is transformed to a spherical grid of the corresponding radar. For example, simulated Mira-35 radar data is transformed to a spherical grid with a range resolution of about 31 m and a maximum range of 24 km. The transformation utilizes the source code radar_filter, which is available on the website of the Stonybrook University together with the CR-SIM source code (https://you.stonybrook.edu/radar/research/radar-simulators/; last accessed 21.09.2021). The radar_filter considers beam propagation effects. I.e., for the interpolation to a grid point of the target spherical grid, all Cartesian input grid points
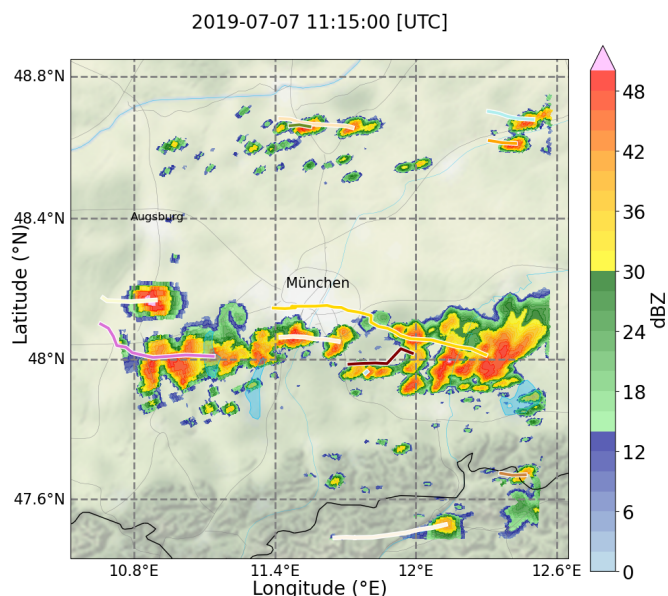
Figure 2. Example of cell-tracking with TINT: Colored background is the reflectivity simulated with WRF and CR-SIM, solid lines and numbers represent the TINT tracks and TINT cell identifier.

255 that are within the beam width are included. If no Cartesian grid point falls into the radar beam, the nearest grid point is used. In the next step attenuation correction is applied along the beam. The correction is applied by subtracting the cumulated (along the range coordinate) simulated attenuation from the uncorrected reflectivity

$$Z_{\mathrm{corr,r}} = Z_{\mathrm{r}} - 2 \cdot \Delta r \cdot \sum_{i=0}^{i=r} A_{\mathrm{i}}. \tag{4}$$

Here the simulated reflectivity without attenuation correction at range gate r is given by Z_{r}. A_{i} is the simulated attenuation
260 in db/m at range gate i, and Δr is the radar range resolution in m. The factor 2 takes into account the fact that the beam travels twice through each grid box (from antenna to target and back). In the same way, the differential reflectivity Z_{DR} is corrected with the simulated differential attenuation ADP.

In a last step, all data (model and radar) is transformed back to a Cartesian grid that exactly covers the Munich domain of the model (144 km times 144 km) with a 400 m times 400 m times 100 m resolution (horizontal x times horizontal y times
265 vertical resolution). This is done by applying a nearest neighbor interpolation that chooses the closest radar bin for each of the Cartesian grid point. Only grid boxes within the lowest and highest radar beam are considered. All grid boxes below the lowest or above the highest beam are masked out. Then, the cell-tracking with TINT is applied to this Cartesian grid in exactly the same way for model and radar data, by passing Py-ART grid objects (Helmus and Collis, 2016) created from the Cartesian grid data to TINT.

Atmospheric
Measurement
Techniques
Discussions

Open Access

EGU
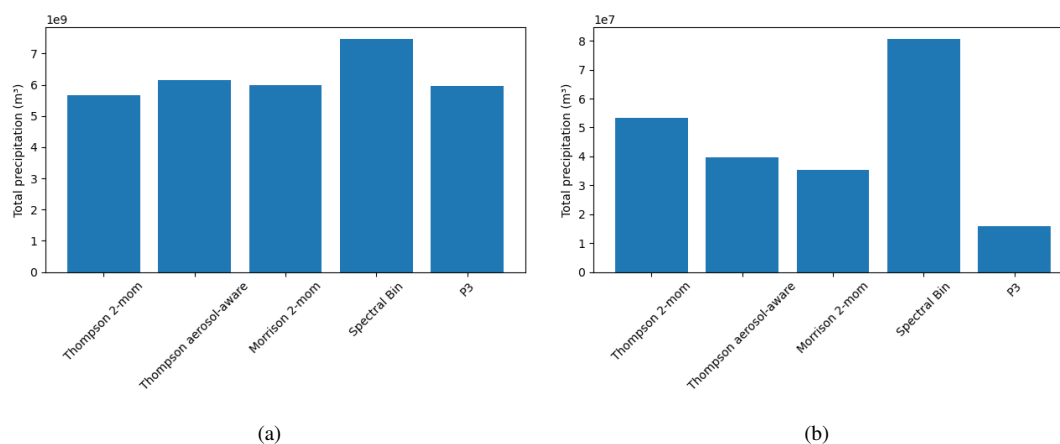
Figure 3. Total simulated precipitation over a) 30 days and b) on 01.08.2020 over the Munich area.

3   Comparison of model and radar observations

An example for the impact of the microphysics scheme choice is given in Figure 3. It shows the total accumulated precipitation over the Munich area as simulated by WRF simulations while only varying the microphysics scheme. Overall, the total accumulated precipitation over the whole domain and over longer periods (3a; 30 days) is similar between all schemes except the bin scheme. However, the deviations can be much larger during single days (e.g., 3b; 1st of August 2020). The total precipitation over all surface grids varies between all 5 schemes, in this case by more than $6 \cdot 10^7 \, \mathrm{m^3}$ between the P3 and the bin scheme. This illustrates the variation between simulations as a result of the choice of microphysics scheme alone. In the following part, we analyze the resulting deviations in more detail.

3.1   Cloud geometry and frequency

We begin our comparison with an evaluation of the geometric properties of simulated and observed clouds. Figure 4 shows histograms of the convective cell extend (area and altitude) as well as the maximum cell reflectivity provided by strategy A. At each 5-min time step during 30 convective weather days all cell detections are summed up on DWD Isen observation data or CR-SIM forward simulations. This means, this analysis is independent of possible matching errors of the cell-tracking, as the identified convective cells at each time step are counted independently. E.g., a single cell detected for 30 min would contribute to the statistics six times. The 32 dBZ cloud top heights of observed cells (Fig. 4a) show a distinct peak with more than 2000 cell detections at an altitude of 3-4 km. This corresponds to about 40 % of all cell detections by the radar. All microphysics schemes in NWP simulations are able to reproduce a peak at a similar altitude but none of them as pronounced as in the observations. The two Thompson schemes show a tendency towards slightly higher cloud tops of 4-5 km. Reflectivities of more than 32 dBZ above the melting layer are mostly related to big graupel particles in our simulations, and to a lesser extent

Atmospheric
Measurement
Techniques
Discussions
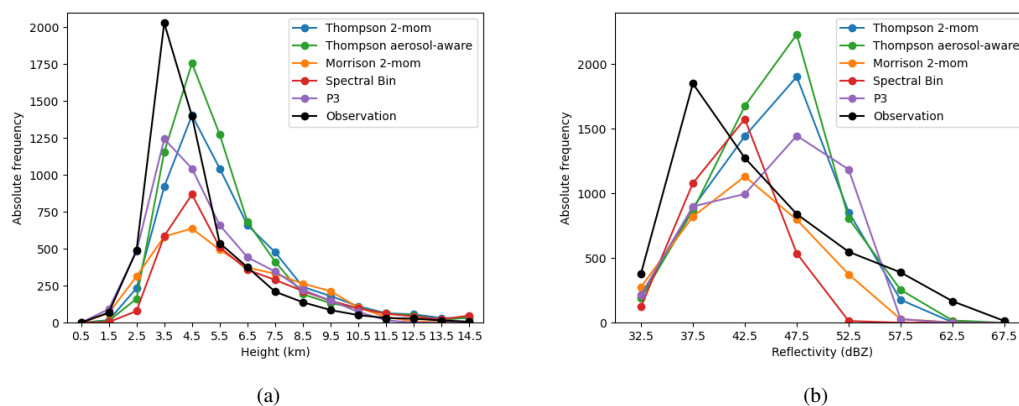
Figure 4. Cell cloud top height (a) and cell maximum reflectivity (b) distribution for inner simulation domain over 30 convection days for observations and 5 microphysics scheme simulations.

rain and snow likely lifted by updrafts. Especially the Thompson schemes more frequently simulate large graupel particles that
290   produce very high reflectivities of more than 45 dBZ above the melting layer (not shown).

Regarding the total number of cell detections the Thompson schemes are closest to the observed number. 5458 cell detections are counted, i.e., the number of cells in all 5-minute observation time steps. The Thompson-aerosol aware scheme (6035) is still close to the observed number; the basic Thompson scheme (5468) is the closest; and the P3 (4768) has fewer cells. Especially the Morrison- (3427) and FSBM-scheme (3326) produce too few convective cells. This difference is mainly a result of missing
295   small-scale development (early stages, weak cells) in the simulations. For fully developed thunderstorms (cell tops > 7 km) all schemes produce numbers that slightly larger than in the observations (observations: 554, Thompson: 1139, Thompson aerosol-aware: 948, Morrison: 928, FSBM: 899, P3: 778). The related distribution of maximum reflectivity of each cell provides some clarification (Fig. 4b). The observed high occurrence of weaker cells is only partially visible in FSBM and Morrison schemes. While the total number of weaker cells (max cell reflectivity at 35 - 40 dBZ) is still too low, the Morrison and FSBM schemes
300   show the highest relative occurrences for relatively weak cells between 40 - 45 dBZ maximum reflectivity. This does still not represent the pronounced peak of observed cells at weaker reflectivities of 35 - 40 dBZ well. The other three schemes produce too many medium intensity cells and too few low intensity cells. At the other end of the reflectivity spectrum, none of the models is able to reproduce the occurrence of the strongest reflectivities at more than 57 dBZ. In part, this is most likely related to numerical smoothing of local and rare values in the NWP model.
305   A similar approach to compare cloud geometry in simulation and radar observation was followed in Caine et al. (2013). They objectively compare simulated cell characteristics with observations over 4.5 days after applying a cell-tracking algorithm on their data. Apart from a larger statistics and better model resolution in the present study, our measurements provide targeted dual-wavelength and polarimetric observations that have the potential to not only provide statistics of general convective cell characteristics, but also to provide vertical profiles of polarimetric and dual-wavelength variables through convective clouds.
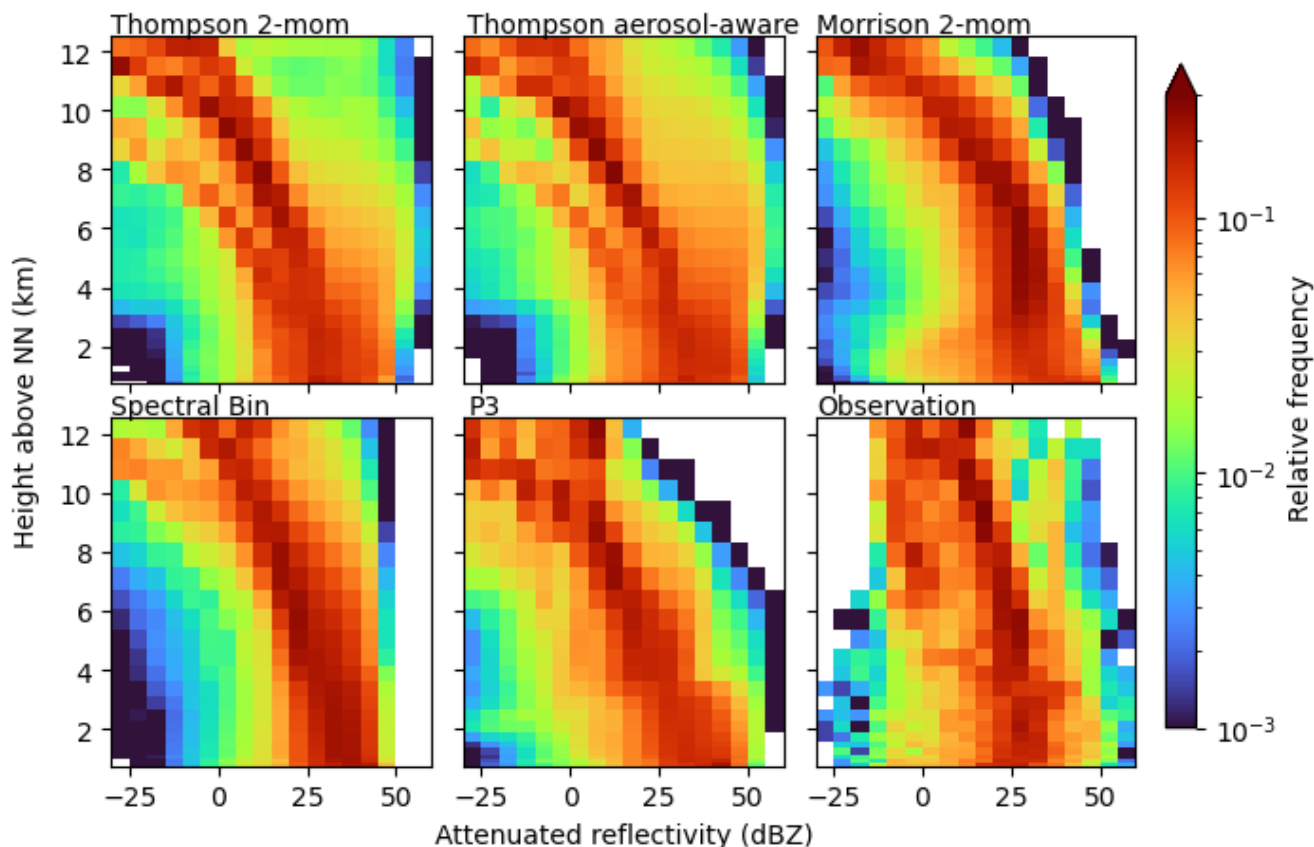
Figure 5. CFADs of simulated and measured reflectivity over 5 convective days in 2019. Radar observations with Poldirad.

310 This has the potential for microphysical fingerprinting in observations as well as simulations. Hence it will help to pinpoint microphysical reasons for uncertainties connected to cloud microphysics which is presented in the next part.

3.2 Profiles of Reflectivity

Contoured frequency by altitude distributions (CFAD) for reflectivity of observed and simulated convective cells are shown in Fig. 5 provided by scan strategy B. This scan strategy provides dual-frequency profiles of high vertical resolution through

315 convective clouds. The radar observation CFADs contain about 1300 profiles in convective clouds. The simulated CFADs consist of many more profiles (on the order of 10^5), because (1) all cells present during one time step on the model domain are analyzed and (2) all columns within each identified cell are included (opposed to the 3x3 profiles that an S-RHI observation provides). The restriction to the center profile of the convective cell, which is a default output of the TINT cell-tracking, would have been an alternative approach. We decided against it for three reasons: (1) the observation was targeted at the location

320 of highest reflectivity and the geometric TINT cell center is not necessarily the location of highest reflectivity; (2) using the

S-RHI strategy we include more variation from each cell compared to one center profile; (3) more profiles provide a better statistical basis for intercomparison of schemes. In the image 5 the simulated reflectivities are corrected for attenuation to make them comparable to the radar observations. Below the melting layer high reflectivities of more than 30 dBZ up to 45 dBZ are simulated most frequently. Overall the schemes agree in the simulated reflectivity in this area mostly caused by rain.

325 It differs only in the spread. The Morrison scheme shows a higher spread, more often simulating reflectivities below 30 dBZ and even down to 0 dBZ. In contrast, the FSBM produces reflectivities below 25 dBZ less often than the others within the convective cells. Compared to the observed CFAD, the simulated reflectivities below the melting layer are generally a little too high.

Above the melting layer simulated reflectivities start to decrease with height. This is a fingerprint of ice growth processes
330 where falling particles increase in size by deposition, aggregation or riming. At these subfreezing heights the schemes show more deviations from each other. While most schemes exhibit a smooth transition from ice to liquid phase, the prominent exception is the P3 scheme for which reflectivities abruptly increase by about 15 dBZ at the melting layer height. All other schemes show a slow and smooth increase in reflectivity, which better agrees with our observations. However, given that the reflectivity within rain was too high, the reflectivity distribution above the melting layer height is reproduced quite well by
335 the P3 scheme. All other schemes simulate higher frequencies of more than 25 dBZ too often. The Thompson schemes even simulate reflectivities of more than 45 dBZ above the melting layer height frequently. These extreme reflectivity values are produced by very large graupel (not shown). Compared to our measurements these reflectivities are unrealistically large. This points to graupel particles that are too large, especially in FSBM and most extreme in the Thompson schemes. At the same time Morrison and FSBM only rarely produce reflectivities of less than 15 dBZ at these heights. The reflectivity overestimation in
340 deep convection at subfreezing temperatures was found by other studies as well (e.g., Stanford et al., 2017; Varble et al., 2011) and is explained to be a result of too large graupel or snow particles, likely a product of overly strong updrafts. Stanford et al. (2017) show that this bias not only exists for bulk schemes, but also for a bin scheme. We can confirm that with our simulations: The bias exists for the FSBM scheme too. Note that the differences seen in the CFADs are only a result of the choice of the microphysics scheme, as the simulations are identical in all other regards. However, as mentioned in the introduction, we are
345 not able to separate the original microphysical impact from other influences, which means that the microphysics schemes are not necessarily the reason alone for the differences seen and feedbacks with the dynamics could also influence the analysis.

3.3 Profiles of polarimetric variables

The same analysis is possible for simulated and observed polarimetric variables, e.g., differential reflectivity Z_{DR} (Fig. 6). Strong differences between the simulations are visible in the liquid phase below the melting layer. While most schemes show
350 a wide spread over the whole range of 0-4 dB within rain, the FSBM only produces Z_{DR} values up to around 1.5 dB. This is in much better agreement with the observations where Z_{DR} values of up to 1.5 have been measured most of the time, though also covering slightly higher Z_{DR}. Here, the advantage of the FSBM that uses a discrete PSD becomes apparent. All other schemes use a gamma distribution (Eq. 1) with a shape parameter $\mu = 0$ for rain. This effectively is an exponential (Marshall-Palmer) PSD which has a slope that is too weak: There are too few small rain droplets and too many big droplets. The FSBM model on

Atmospheric
Measurement
Techniques
Discussions

Open Access

EGU
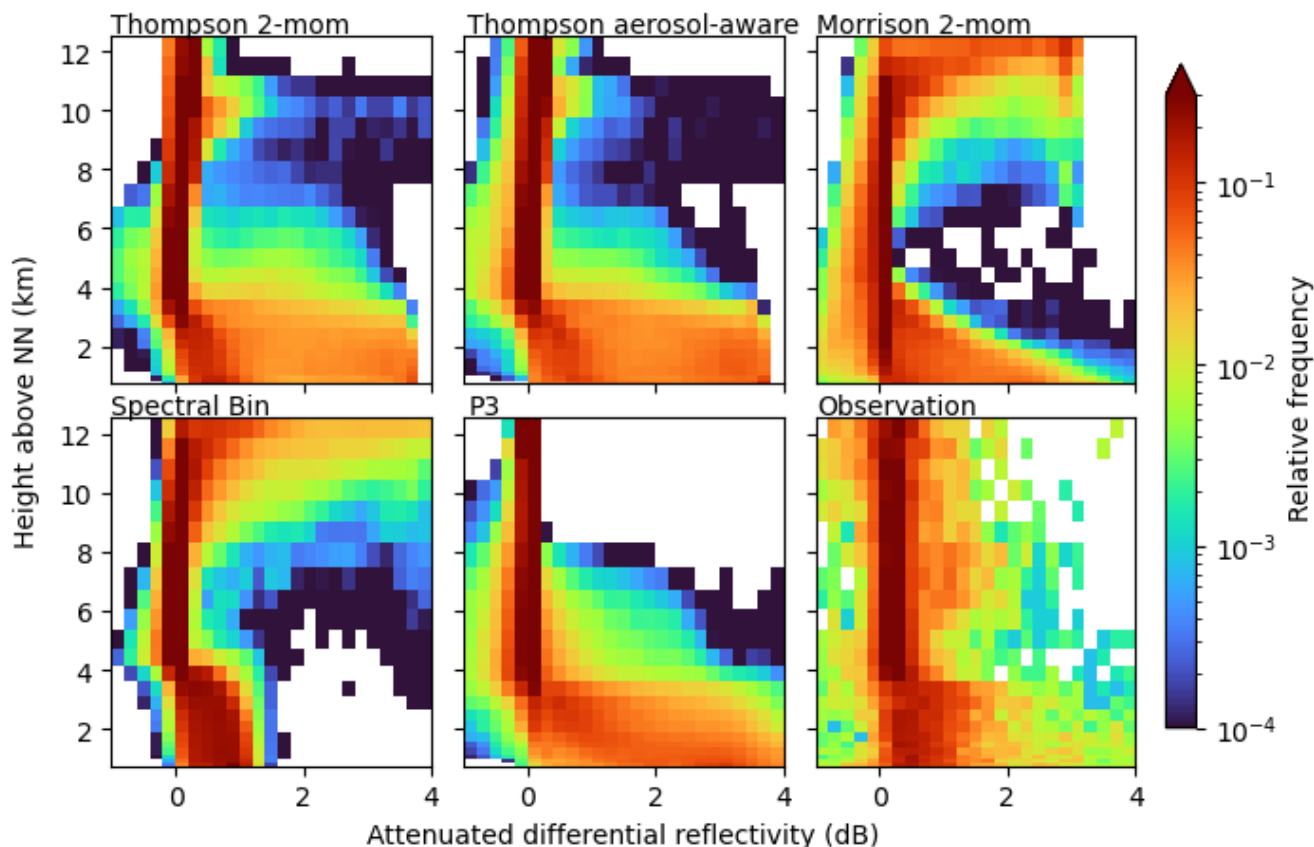
Figure 6. CFADs of simulated and measured differential reflectivity over 5 convective days in 2019. Observation with the Poldirad radar.

355 the other hand is able to explicitly predict rain droplets of each bin which results in a more realistic PSD (better size-sorting). Hence, radar Z_{DR} signatures are closer to our observations.

Directly above the melting layer the FSBM and Morrison schemes show Z_{DR} values close to 0, while the P3 and the Thompson schemes have their frequency maximum at 0, but show more spread also to higher Z_{DR} values. Z_{DR} of 0 is associated with spherical particles. This reflects the spherical nature of the simulated idealized ice particles in the Morrison and FSBM scheme.

360 In the Thompson schemes snow is assumed to be oblate (mass-size relation parameter $b = 2$), and in the P3 scheme partially rimed and unrimed large ice particles are assumed to be more flat (mass-size relation parameter $b = 1.9$). This results in Z_{DR} values greater than 0 and produces more spread in the CFAD which is more realistic and fits our measurements better.

### 3.4 Profiles of Dual-Wavelength variables

More insight about the particle size is provided by the simulated and observed dual-wavelength ratio DWR. The standard
365 radar reflectivity is strongly influenced by the number of particles within the radar beam, the particles sizes and the particles
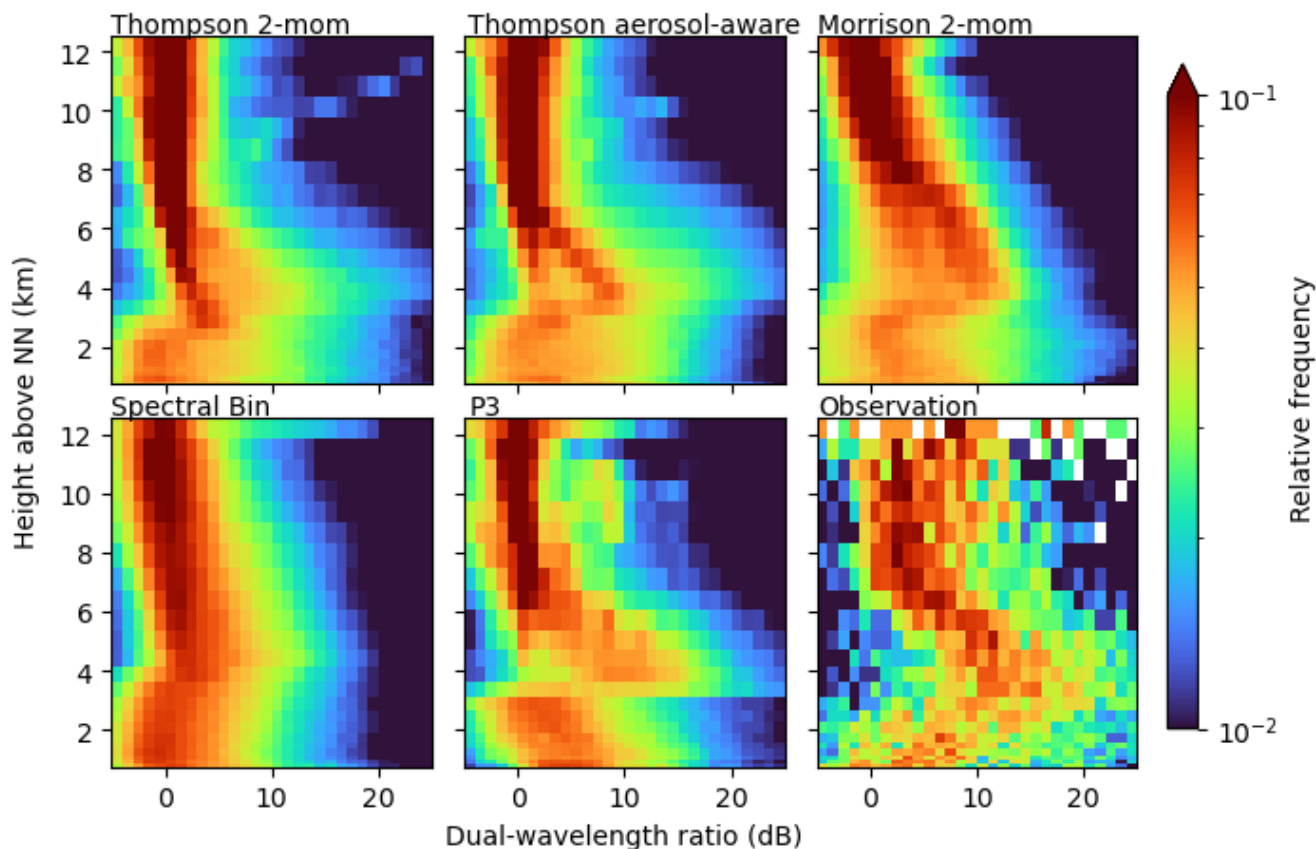
Figure 7. CFADs of simulated and measured DWR over 5 convective days in 2019. Radar observations with Poldirad and Mira35 (Poldirad - Mira35).

densities. In contrast, DWR is influenced not so much by the density or the number of particles but is rather sensitive to the particles size. Figure 7 shows deviations between the schemes within the ice phase as well as in the liquid phase. Here, no attenuation correction is applied. This makes the comparison to the radar observations less realistic but reveals differences in microphysical processes and fingerprints between the simulations more clearly. The observations show DWR close to 0 at greater heights where ice crystals are very small. All simulations agree at these heights. The observations then show a steady increase of DWR towards the melting layer height. This is reflecting ice particle growth, given that DWR is mainly sensitive to particle size. All simulations reproduce this increase of DWR towards the melting layer heights but differ in the slope and height where the increase starts. While the Morrison and FSBM simulations show a beginning increase in DWR already at about 10 km, the P3 and the two Thompson simulations show a beginning increase at about 7 km, which better agrees with the observations. At melting layer heights the DWR values reach their maximum in all simulations. The magnitude of the maximum DWR values differs: Morrison and FSBM do not produce DWR larger than 20 dB, while the P3 and the

370

375

Atmospheric
Measurement
Techniques
Discussions
Open Access
EGU
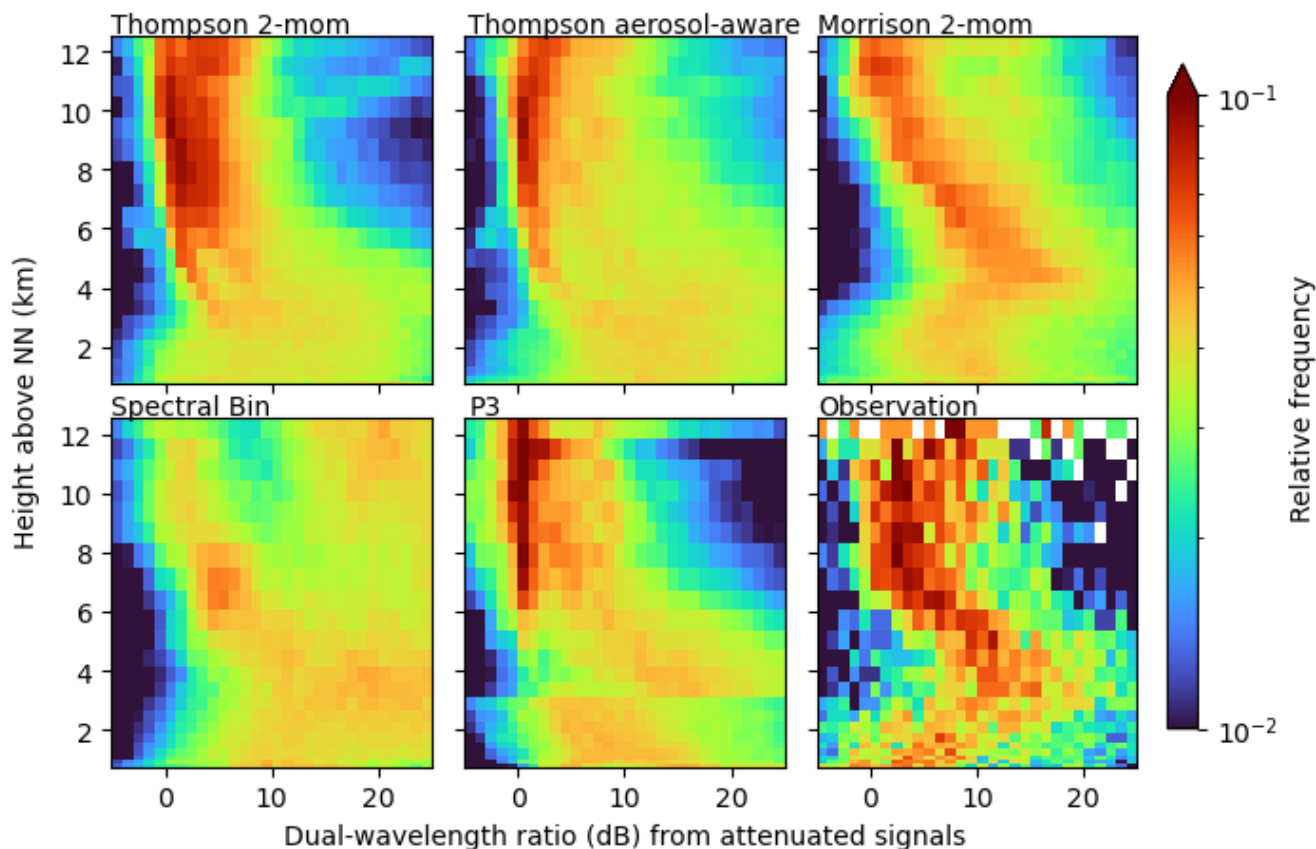
Figure 8. Same as Fig. 7 but with simulated attenuation included in the simulations.

two Thompson schemes produce DWR of up to 25 dB. At these heights the two Thompson schemes produce corridors of higher frequencies at lower DWR values (0-10 dB) and then a diffuse area of lower frequencies at higher DWR (> 10 dB). The corridors are likely related to snow growth during sedimentation. Thompson only uses one constant mass-size relation for snow with a constant density that is independent of its size. The P3 is more flexible: It uses a varying mass-size relation depending

380 on whether the ice particle is unrimed, partially rimed or fully rimed. This might be the reason why the DWR corridor in P3 above the melting layer height is wider. The high DWR values of more than 20 dB above the melting layer is caused by graupel particles (not shown). Graupel particles are much denser than snow and therefore have more mass and produce much higher reflectivities than snow at the same size. These diffuse areas of high DWR values in the Thompson schemes are related to the

385 extreme high values seen in the reflectivity CFADs (Fig. 5).

Below the melting layer the observed DWR steadily decreases towards the ground. The models do not reproduce this very well: Even though the DWR decreases in all models, this decrease happens abruptly at the melting layer. Below this height the DWR stays more or less constant. In the P3 simulations (and weaker in the Morrison scheme) the DWR even increases again

Atmospheric
Measurement
Techniques
Discussions

Open Access
EGU

towards the ground. At these heights rain is the dominant species. The simulated increase of DWR towards the ground is likely

390 a result of the simulated collection process: Rain droplets grow while falling by collecting smaller droplets. This process is not visible in the observed DWR. This could be an artifact of the smaller number of profiles that are included in the observations. However, it should be emphasized that the DWR signatures presented here are very hard to observe with vertical pointing radars due to problems arising from a wet radome or antenna and very quick attenuation of the signal below a convective cell. This demonstrates one advantage of our scanning setup, where we can measure all layers of convective clouds without being

395 directly below the convective cloud. The general magnitude of DWR near the surface was measured at around -3 to 10 dB in observations as well as all simulations. Note that including attenuation will increase these values, as the Ka-band reflectivity is attenuated more strongly and hence the DWR increases.

Comparing DWR signatures without including attenuation is giving insight into the details of the microphysical schemes but is not well suited for a direct comparison with radar observations, because especially the Ka-band observations are potentially

400 strongly attenuated. Figure 8 shows the same DWR CFADs including attenuation. Obviously attenuation drastically increases the variability in DWR. As a result, DWR values are scattered over larger ranges, partially masking the underlying fingerprints that were visible in the CFADs without attenuation. Furthermore, the lower number of observed profiles compared to the simulated profiles is also clearly visible. This is most prominent in the DWR CFAD but also in the CFADs of reflectivity (Fig. 5) and Z_{DR} (Fig. 6). This is reminiscent of the large observational effort to collect targeted cell core RHI scans.

405 ## 4   Summary and conclusions

A methodological framework has been presented that allows for a statistical comparison of polarimetric dual-wavelength radar observations with numerical weather model output. Targeted dual-radar observations of convective cell characteristics in the vicinity of Munich over a significant fraction of cell lifetime have been established. For the weather model, cell specific observations of a Ka-band and two C-band radars are produced using a polarimetric radar forward operator and automatic

410 cell-tracking. The total data set presented includes 30 convective days of simulation and radar observations so far. Targeted dual-wavelength observations were performed on 5 of those days, adding up to about 1300 RHI profiles of dual-wavelength observations of convective clouds. A convection permitting Weather Research and Forecast Model (WRF) setup over Munich has been implemented. WRF hindcast simulations were conducted with 5 different microphysics schemes of varying complexity. The Cloud Resolving Model and Radar Simulator (CR-SIM) was applied that provides polarimetric radar variables from

415 model output. The cell-tracking algorithm TINT is not TITAN (TINT) was applied on radar and model data in the same way to allow for comparison of convective cell characteristics.

With the exception of the Thompson schemes, all microphysics schemes simulate too few convective cells compared to the radar observations. The difference is mainly caused by missing weak cells with cloud top heights between 3-4 km. This points to missing small scale development in most of the simulations. This suggests dynamical reasons rather than impact of

420 microphysics schemes. It is a reminder that the presented methodology is not able to perfectly separate microphysics from dynamical feedbacks. Statistics of observed and simulated cell reflectivities show that models are not able to reproduce the

Atmospheric
Measurement
Techniques
Discussions
Open Access

EGU

observed high occurrences of very weak cells, as well as the occurrence of strongest reflectivities at more than 57 dBZ. This might be related to numerical smoothing of local and rare values in the Numerical Weather Prediction (NWP) model.

Targeted scans of convective clouds revealed differences in radar observations and model as well as between microphysics
425   schemes in ice and liquid phases. Overall the schemes agree in simulated reflectivity in the liquid phase, but comparison to the radar observations show reflectivity is slightly overestimated in all schemes. Simulated Z_{DR} reveal that only the Fast Spectral Bin Microphysics (FSBM) scheme is able to reproduce these radar signatures reasonably well, all other schemes produce Z_{DR} signals within rain with too much spread. This is likely a result of the assumed exponential (Marshall-Palmer) rain particle size distribution (PSD) producing too many large and too few small droplets. The FSBM scheme on the other hand demonstrates
430   the advantage of explicitly resolving the PSD that results in a more realistic radar signature in rain.

Within the ice phase more deviations between the schemes are found. FSBM and Morrison simulations were not able to reproduce the high frequency of observed small reflectivities at 4-6 km altitude. The Predicted Particle Property (P3) scheme is the only scheme not overestimating reflectivities above the melting layer height. All other schemes show unrealistic high reflectivities, partly over 45 dBZ, which mainly is a result of the presence of too many large graupel particles. This ice size
435   bias is likely related to wrong graupel and snow particle size distributions and was found in other studies as well (e.g., Stanford et al., 2017; Varble et al., 2011). Contoured Frequency by Altitude Distributions (CFADs) of dual-wavelength ratio (DWR) allow to analyze the simulated particles sizes more directly. The two Thompson schemes produce clearly confined distributions of higher occurrences of DWR in the ice phase related to snow and its growth by aggregation during sedimentation. The P3 scheme also produces distinct narrow distribution of DWR values at these heights. Nonetheless it is wider compared to the
440   Thompson results, caused by a mixture of unrimed, partly rimed and unrimed particles. We believe this demonstrates the greater flexibility of the P3 scheme in the ice phase where this scheme deviates from the traditional hydrometeor classes and instead predicts properties for a single ice category. This seems to produce a wider range of particle characteristics (and hence DWR signals) as opposed to the other schemes, where most of the signal is produced along distinctly visible corridors. CFADs of Z_{DR} support this impression: Only the P3 and the Thompson schemes reproduce the observed spread of Z_{DR} above the
445   melting layer. All other schemes simulate Z_{DR} values close to zero, revealing the idealized spherical nature of the simulated ice particles.

In general, we could demonstrate how weather simulations with varying microphysics schemes produce varying polarimetric and DWR radar signatures. However, one interesting fact is that the two Thompson schemes do not show significant differences from one another. Even though the schemes are very similar, one could have expected that the explicit prediction of droplet
450   number concentration as well of aerosol variables would have a stronger influence on the weather simulation.

The analyses shown in this work demonstrate the potential to analyze the treatment of small scale processes within microphysics schemes. More analyses will be conducted with the methods presented, especially including dual-wavelength and polarimetric variables to analyze the simulated particle shapes and sizes. The observed radar CFADs still show large scatter due to small numbers of measurements included. More dual-wavelength data is needed to compare radar observations and a
455   model for convective weather situations with more confidence. Another operational dual-wavelength measurement strategy is currently being established that makes use of the operational DWD volume scans and copies their strategy with the Mira-35

Ka-band radar. Because the volume scan strategy consists of multiple Plan Position Indicator (PPI) scans of different eleva-tions angles, the vertical resolution will be somewhat lower compared to our dual-wavelength Range Height Indicator (RHI) scans in strategy B. On the other hand, the PPI strategy possibly samples multiple cells at the same time and together with the

460 operational setup we expect to obtain a larger number of dual-wavelength measurements of convective cells.

Based on the methodology presented in this paper, more detailed analysis of some of the observed differences will be analysed next. This will allow to slowly approach the answer to the question which level of complexity in microphysical processes needs to be implemented to realistically represent cloud and precipitation distribution and radiation budgets in NWP models at the same time.

465 *Code and data availability.* The polarimetric radar data from the operational C-band radar in Isen is available for research from the German Weather Service (DWD) upon request. The Poldirad and Mira-35 data presented in this paper is available through the au-thors upon request. Data of WRF, CR-SIM, and TINT simulations is also available through the authors upon request. The software developed for this paper is available here: doi.org/10.5281/zenodo.5526882. The Weather Research and Forecasting Model (WRF; ver-sion 4.2) is openly available on github: https://github.com/wrf-model/WRF. The cell-tracking algorithm TINT is openly available on github:

470 https://github.com/openradar/TINT (last accessed 21.09.2021). The forward operator CR-SIM (version 3.33) is available at the website of the Stonybrook University (https://you.stonybrook.edu/radar/research/radar-simulators/; last accessed 21.09.2021).

Appendix A: Simulation and Observation dates

Table A1. List of convective days that were used in our analyses. Strategy A always refers to the whole day.

Date	Strategy
29.04.2019	Strategy A
06.05.2019	Strategy A
28.05.2019	Strategy A, Strategy B (11:25 - 14:00 UTC)
29.05.2019	Strategy A
11.06.2019	Strategy A
12.06.2019	Strategy A
21.06.2019	Strategy A, Strategy B (14:40 - 17:25 UTC)
01.07.2019	Strategy A, Strategy B (11:20 - 16:50 UTC)
07.07.2019	Strategy A, Strategy B (09:20 - 15:10 UTC)
08.07.2019	Strategy A, Strategy B (09:00 - 14:00 UTC)
17.06.2020	Strategy A
20.06.2020	Strategy A
27.06.2020	Strategy A
28.06.2020	Strategy A
29.06.2020	Strategy A
01.07.2020	Strategy A
10.07.2020	Strategy A
11.07.2020	Strategy A
23.07.2020	Strategy A
24.07.2020	Strategy A
26.07.2020	Strategy A
28.07.2020	Strategy A
01.08.2020	Strategy A
02.08.2020	Strategy A
03.08.2020	Strategy A
18.08.2020	Strategy A
17.09.2020	Strategy A
22.09.2020	Strategy A
23.09.2020	Strategy A
12.10.2020	Strategy A

Atmospheric
Measurement
Techniques
Discussions

Open Access

EGU

Author contributions. Eleni Tetoni, Florian Ewald, Martin Hagen and Gregor Köcher performed radar measurements during precipitation events. Gregor Köcher developed the methodology presented and wrote the manuscript in its current form. Tobias Zinner and Christoph
475 Knote supervised and discussed the scientific content. All authors commented on the manuscript.

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

Atmospheric
Measurement
Techniques
Discussions
Open Access
EGU

References

Baldauf, M., Seifert, A., Förstner, J., Majewski, D., Raschendorfer, M., and Reinhardt, T.: Operational convective-scale numerical weather prediction with the COSMO model: Description and sensitivities, Monthly Weather Review, 139, 3887–3905, 2011.

485 Bauer-Pfundstein, M. R. and Görsdorf, U.: Target separation and classification using cloud radar Doppler-spectra, in: Proceedings 33rd Intern. Conf. on Radar Meteorology, Cairns, Australia, 2007.

Brown, P. R. and Francis, P. N.: Improved measurements of the ice water content in cirrus using a total-water probe, Journal of Atmospheric and Oceanic Technology, 12, 410–414, 1995.

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

490 model simulations using a cell-tracking algorithm, Monthly Weather Review, 141, 557–581, 2013.

Cao, Q., Zhang, G., Brandes, E., Schuur, T., Ryzhkov, A., and Ikeda, K.: Analysis of video disdrometer and polarimetric radar data to characterize rain microphysics in Oklahoma, Journal of Applied Meteorology and Climatology, 47, 2238–2255, 2008.

Dixon, M. and Wiener, G.: TITAN: Thunderstorm identification, tracking, analysis, and nowcasting—A radar-based methodology, Journal of atmospheric and oceanic technology, 10, 785–797, 1993.

495 Fan, J., Liu, Y.-C., Xu, K.-M., North, K., Collis, S., Dong, X., Zhang, G. J., Chen, Q., Kollias, P., and Ghan, S. J.: Improving representation of convective transport for scale-aware parameterization: 1. Convection and cloud properties simulated with spectral bin and bulk microphysics, Journal of Geophysical Research: Atmospheres, 120, 3485–3509, 2015.

Fan, J., Han, B., Varble, A., Morrison, H., North, K., Kollias, P., Chen, B., Dong, X., Giangrande, S. E., Khain, A., et al.: Cloud-resolving model intercomparison of an MC3E squall line case: Part I—Convective updrafts, Journal of Geophysical Research: Atmospheres, 122,

500 9351–9378, 2017.

Field, P., Hogan, R., Brown, P., Illingworth, A., Choularton, T., and Cotton, R.: Parametrization of ice-particle size distributions for mid-latitude stratiform cloud, Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography, 131, 1997–2017, 2005.

Flack, D. L., Gray, S. L., and Plant, R. S.: A simple ensemble approach for more robust process-based sensitivity analysis of case studies in

505 convection-permitting models, Quarterly Journal of the Royal Meteorological Society, 145, 3089–3101, 2019.

Fridlind, A., van Lier-Walqui, M., Collis, S., Giangrande, S., Li, X., Matsui, T., Orville, R., Rosenfeld, D., Weitz, R., and Zhang, P.: Use of polarimetric radar measurements to constrain simulated convective cell evolution: A pilot study with Lagrangian tracking., Atmospheric Measurement Techniques, 12, 2019.

Gallus Jr, W. A. and Pfeifer, M.: Intercomparison of simulations using 5 WRF microphysical schemes with dual-polarization data for a

510 German squall line, Advances in Geosciences, 16, 109, 2008.

Grabowski, W. W.: Extracting microphysical impacts in large-eddy simulations of shallow convection, Journal of the Atmospheric Sciences, 71, 4493–4499, 2014.

Han, B., Fan, J., Varble, A., Morrison, H., Williams, C. R., Chen, B., Dong, X., Giangrande, S. E., Khain, A., Mansell, E., et al.: Cloud-resolving model intercomparison of an MC3E squall line case: Part II. Stratiform precipitation properties, Journal of Geophysical Re-

515 search: Atmospheres, 124, 1090–1117, 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, DWD and DLR, vol. 4, 2014.

Helmus, J. J. and Collis, S. M.: The Python ARM Radar Toolkit (Py-ART), a library for working with weather radar data in the Python
520   programming language, Journal of Open Research Software, 4, 2016.

Heymsfield, A. J.: Properties of tropical and midlatitude ice cloud particle ensembles. Part II: Applications for mesoscale and climate models,
       Journal of the atmospheric sciences, 60, 2592–2611, 2003.

Jankov, I., Grasso, L. D., Sengupta, M., Neiman, P. J., Zupanski, D., Zupanski, M., Lindsey, D., Hillger, D. W., Birkenheuer, D. L., Brummer,
       R., et al.: An evaluation of five ARW-WRF microphysics schemes using synthetic GOES imagery for an atmospheric river event affecting
525   the California coast, Journal of Hydrometeorology, 12, 618–633, 2011.

Johnson, J., Cui, Z., Lee, L., Gosling, J., Blyth, A., and Carslaw, K.: Evaluating uncertainty in convective cloud microphysics using statistical
       emulation, Journal of Advances in Modeling Earth Systems, 7, 162–187, 2015.

Kessler, E.: On the distribution and continuity of water substance in atmospheric circulations, in: On the distribution and continuity of water
       substance in atmospheric circulations, pp. 1–84, Springer, 1969.

530   Khain, A., Pokrovsky, A., Pinsky, M., Seifert, A., and Phillips, V.: Simulation of effects of atmospheric aerosols on deep turbulent convective
       clouds using a spectral microphysics mixed-phase cumulus cloud model. Part I: Model description and possible applications, Journal of
       the atmospheric sciences, 61, 2963–2982, 2004.

Khain, A., Beheng, K., Heymsfield, A., Korolev, A., Krichak, S., Levin, Z., Pinsky, M., Phillips, V., Prabhakaran, T., Teller, A., et al.:
       Representation of microphysical processes in cloud-resolving models: Spectral (bin) microphysics versus bulk parameterization, Reviews
535   of Geophysics, 53, 247–322, 2015.

Lean, H. W., Clark, P. A., Dixon, M., Roberts, N. M., Fitch, A., Forbes, R., and Halliwell, C.: Characteristics of high-resolution versions of
       the Met Office Unified Model for forecasting convection over the United Kingdom, Monthly Weather Review, 136, 3408–3424, 2008.

Li, X., Tao, W.-K., Khain, A. P., Simpson, J., and Johnson, D. E.: Sensitivity of a cloud-resolving model to bulk and explicit bin microphysical
       schemes. Part I: Comparisons, Journal of the Atmospheric Sciences, 66, 3–21, 2009.

540   Li, Z., Zuidema, P., Zhu, P., and Morrison, H.: The sensitivity of simulated shallow cumulus convection and cold pools to microphysics,
       Journal of the Atmospheric Sciences, 72, 3340–3355, 2015.

Martin, G., Johnson, D., and Spice, A.: The measurement and parameterization of effective radius of droplets in warm stratocumulus clouds,
       Journal of Atmospheric Sciences, 51, 1823–1842, 1994.

Milbrandt, J. and Yau, M.: A multimoment bulk microphysics parameterization. Part II: A proposed three-moment closure and scheme
545   description, Journal of Atmospheric Sciences, 62, 3065–3081, 2005.

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, 2015.

Morrison, H. and Pinto, J.: Intercomparison of bulk cloud microphysics schemes in mesoscale simulations of springtime Arctic mixed-phase
       stratiform clouds, Monthly weather review, 134, 1880–1900, 2006.

550   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, 2009.

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., et al.: Confronting the challenge of modeling cloud and precipitation microphysics, Journal of Advances in Modeling
       Earth Systems, p. e2019MS001689, 2020.

555   NCEP: 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.: The Cloud-resolving model Radar SIMulator (CR-SIM) Version 3.3: description and applications of a virtual observatory, Geoscientific Model Development (Print), 13, 2020.

Pinto, J. O., Grim, J. A., and Steiner, M.: Assessment of the High-Resolution Rapid Refresh model's ability to predict mesoscale convective systems using object-based evaluation, Weather and Forecasting, 30, 892–913, 2015.

560 Rajeevan, M., Kesarkar, A., Thampi, S., Rao, T. N., Radhakrishna, B., and Rajasekhar, M.: Sensitivity of WRF cloud microphysics to simulations of a severe thunderstorm event over Southeast India, in: Annales Geophysicae, vol. 28, pp. 603–619, Copernicus GmbH, 2010.

Schroth, A. C., Chandra, M. S., and Mesichner, P. F.: AC-band coherent polarimetric radar for propagation and cloud physics research, Journal of Atmospheric and Oceanic Technology, 5, 803–822, 1988.

565 Seity, Y., Brousseau, P., Malardel, S., Hello, G., Bénard, P., Bouttier, F., Lac, C., and Masson, V.: The AROME-France convective-scale operational model, Monthly Weather Review, 139, 976–991, 2011.

Shima, S.-i., Kusano, K., Kawano, A., Sugiyama, T., and Kawahara, S.: The super-droplet method for the numerical simulation of clouds and precipitation: A particle-based and probabilistic microphysics model coupled with a non-hydrostatic model, Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography, 135, 1307–1320, 570 2009.

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, 2019.

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., and Huang, 575 X.-Y.: A description of the advanced research WRF model version 4, Tech. rep., NCAR/TN-556+STR, 2019.

Stanford, M. W., Varble, A., Zipser, E., Strapp, J. W., Leroy, D., Schwarzenboeck, A., Potts, R., and Protat, A.: A ubiquitous ice size bias in simulations of tropical deep convection, Atmospheric Chemistry and Physics, 17, 9599–9621, 2017.

Stanford, M. W., Morrison, H., Varble, A., Berner, J., Wu, W., McFarquhar, G., and Milbrandt, J.: Sensitivity of simulated deep convection to a stochastic ice microphysics framework, Journal of Advances in Modeling Earth Systems, 11, 3362–3389, 2019.

580 Stein, T. H., 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, 2015.

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, 2014.

585 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, 2008.

Trömel, S., Simmer, C., Blahak, U., Blanke, A., Ewald, F., Frech, M., Gergely, M., Hagen, M., Hörnig, S., Janjic, T., et al.: Overview: Fusion of Radar Polarimetry and Numerical Atmospheric Modelling Towards an Improved Understanding of Cloud and Precipitation Processes, Atmospheric Chemistry and Physics Discussions, pp. 1–36, 2021.

590 Varble, A., Fridlind, A. M., Zipser, E. J., Ackerman, A. S., Chaboureau, J.-P., Fan, J., Hill, A., McFarlane, S. A., Pinty, J.-P., and Shipway, B.: Evaluation of cloud-resolving model intercomparison simulations using TWP-ICE observations: Precipitation and cloud structure, Journal of Geophysical Research: Atmospheres, 116, 2011.

Varble, A., Zipser, E. J., Fridlind, A. M., Zhu, P., Ackerman, A. S., Chaboureau, J.-P., Collis, S., Fan, J., Hill, A., and Shipway, B.: Eval-
uation of cloud-resolving and limited area model intercomparison simulations using TWP-ICE observations: 1. Deep convective updraft
595   properties, Journal of Geophysical Research: Atmospheres, 119, 13–891, 2014.

White, B., Gryspeerdt, E., Stier, P., Morrison, H., Thompson, G., and Kipling, Z.: Uncertainty from the choice of microphysics scheme in
convection-permitting models significantly exceeds aerosol effects, Atmospheric Chemistry and Physics, 17, 12 145–12 175, 2017.

Xue, L., Fan, J., Lebo, Z. J., Wu, W., Morrison, H., Grabowski, W. W., Chu, X., Geresdi, I., North, K., Stenz, R., et al.: Idealized simulations
of a squall line from the MC3E field campaign applying three bin microphysics schemes: Dynamic and thermodynamic structure, Monthly
600   Weather Review, 145, 4789–4812, 2017.