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

Author(s): Gregor Köcher, Tobias Zinner, Christoph Knote, Eleni Tetoni, Florian Ewald, and Martin Hagen

Summary

This study examines a series of regional storm-resolving weather forecasting with different microphysics schemes against long-term observations of dual-wavelength polarimetric radars over Munchen area. Although there are many papers inter-compare microphysics schemes available in WRF, the noble aspect of this study is to utilize polarimetric radar simulator and cell-tracking algorithm for more consistent sampling of polarimetric radar observables and dual-wavelength ratio. However, there are some major questions/suggestions related to 1) missing citations, 2) separation of convective cells, and 3) actual rain drop-size distributions, and 4) uncertainties of the forward model. By improving these issues, this manuscript could be quite powerful. Thus, my suggestion is "major revisions" in order to publish in ACP.

We like to thank the reviewer for the valuable comments which helped to improve the manuscript quality substantially. Please find below our point-by-point reply highlighted in blue. A marked-up manuscript version showing the changes made is provided along with the revised manuscript.

Major Comments/Suggestions

1) Missing citations

This study cites several microphysics evaluation papers, but it completely misses previous studies directly using forward polarimetric radar models. Here are suggested references. Please take a look and relate yours to their findings and approaches.

Jung, Y., M. Xue, and G. Zhang, 2010: Simulations of Polarimetric Radar Signatures of a Supercell Storm Using a Two-Moment Bulk Microphysics Scheme. *J. Appl. Meteor. Climatol.*, **49**, 146–163, https://doi.org/10.1175/2009JAMC2178.1

Snyder, J.C., H.B. Bluestein, D.T. Dawson II, and Y. Jung, 2017a: Simulations of Polarimetric, X-Band Radar Signatures in Supercells. Part I: Description of Experiment and Simulated ρhv Rings. *J. Appl. Meteor. Climatol.*, **56**, 1977–1999, https://doi.org/10.1175/JAMC-D-16-0138.1

Ryzhkov, A., M. Pinsky, A. Pokrovsky, and A. Khain, 2011: Polarimetric Radar Observation Operator for a Cloud Model with Spectral Microphysics. J. Appl. Meteor. Climatol., 50, 873–894. doi: http://dx.doi.org/10.1175/2010JAMC2363.1

Putnam, B.J., M. Xue, Y. Jung, G. Zhang, and F. Kong, 2017: Simulation of Polarimetric Radar Variables from 2013 CAPS Spring Experiment Storm-Scale Ensemble Forecasts and Evaluation of Microphysics Schemes. Mon. Wea. Rev., 145, 49–73, https://doi.org/10.1175/MWR-D-15-0415.1

Also a following citation is recommended for microphysics finger print from polarimetric radar signals. This Ph.D. dissertation describes various microphysics finger prints related to cloud microphysics processes.

Kumjian, M.R., 2012: The impact of precipitation physical processes on the polarimetric radar variables. PhD. Dissertation, University of Oklahoma, 327 pp.

Thank you for your suggestions. We included the proposed literature regarding polarimetric radar forward operators to evaluate cloud model microphysics in our introduction:

> *There are some studies that directly use polarimetric radar forward operators to evaluate the performance of cloud microphysics schemes. For instance, Jung et al. (2010) and Snyder et al. (2017) each simulate idealized super cell events to test if the cloud microphysics schemes together with a polarimetric radar forward operator are able to reproduce known super cell radar signatures. Ryzhkov et al. (2011) and Putnam et al. (2017) compare simulated polarimetric radar signals with radar observations to evaluate microphysics schemes, but focus on one or two convective cases.*

Furthermore, we have related the findings of Ryzhkov et al. (2011) and Putnam et al. (2017) to ours in the discussion. Jung et al. (2010) and Snyder et al. (2017) simulate idealized super cell events, which is a different approach than ours. That's why we think it is enough to relate our approach to theirs in the introduction, but we don't relate our findings to theirs in the discussion.

Putnam et al. (2017), section 3.2:

> *This is in agreement with Putnam et al. (2017) who compare radar signals simulated by 5 different microphysic schemes for two case studies and find that especially the Morrison scheme but to a lesser extent also the Thompson scheme produces too high Z. They attribute this to stratiform rain PSDs that contain too many large drops, to an overforecast of the precipitation coverage overall and in case of Morrison, to a high bias of wet graupel in convective regions. Given that the forward simulator applied in this study does not consider wet particles, we find the high bias in Z exists even without considering wet graupel and comes mostly from rain, suggesting PSDs that contain too many large rain drops compared to the observations.*

Putnam et al. (2017), section 3.3:

> *Putnam et al. (2017) find similar results regarding ZDR signatures near the surface: in their two case studies, the simulations with Thompson and with Morrison cloud microphysics showed incorrect ZDR maxima associated with isolated large drops at locations of weak convection where this would not be expected.*

Ryzhkov et al. (2011), section 3.3:

> *Ryzhkov et al. (2011) for example evaluate radar signals simulated from a spectral bin scheme against a hailstorm case and find that their spectral bin scheme produces PSDs for rain that deviate from the gamma distribution. Bulk schemes would not be able to reproduce these PSDs and since radar signals strongly depend on the PSD, Ryzhkov et al. (2011) argue that spectral bin schemes are better suited to simulate polarimetric radar signals.*

Finally, we referenced the Ph.D. Dissertation about fingerprints in polarimetric radar signals, also in our introduction. This Ph.D. Dissertation helps a lot in understanding the impact of rain processes on polarimetric radar signals:

> *Kumjian (2012) demonstrate the impact of precipitation processes on polarimetric radar signals, though he focuses mainly on rain processes, such as raindrop evaporation or size sorting.*

2) separation of convective cells

One of the advantages of this paper is the large sampling volume, but it simultaneously induces ambiguity for analysis. During a long-term period, there must be various sizes of convections from shallow, congestus, and deep convective/stratiform cells. When you bundle all into a single CFAD, it tends to smear out important aspects of microphysics. Please check the following papers on how it separates cloud type and better evaluates different aspects of microphysics from long-term simulations/observations.

Matsui, T., X. Zeng, W.-K. Tao, H. Masunaga, W. Olson, and S. Lang (2009), Evaluation of long-term cloud-resolving model simulations using satellite radiance observations and multifrequency satellite simulators. *Journal of Atmospheric and Oceanic Technology*, 26, 1261-1274.

You can use echo-top height from each cell-tracked target for separation. But if this type of separation is too difficult to implement (or too much effort), please just discuss and try it in the future.

We have added a part discussing this topic in the conclusions. This is a valuable comment, we plan to include a separation of convection type for the next steps of analysis of the data to be published soon after this manuscript. Apart from a possible separation using the echo-top height for different convection sizes, we think a separation into weak forcing / strong forcing situations could be interesting. We think that this is too much effort for the present work, that’s why we are just discussing it in this manuscript and will add it for future publications.

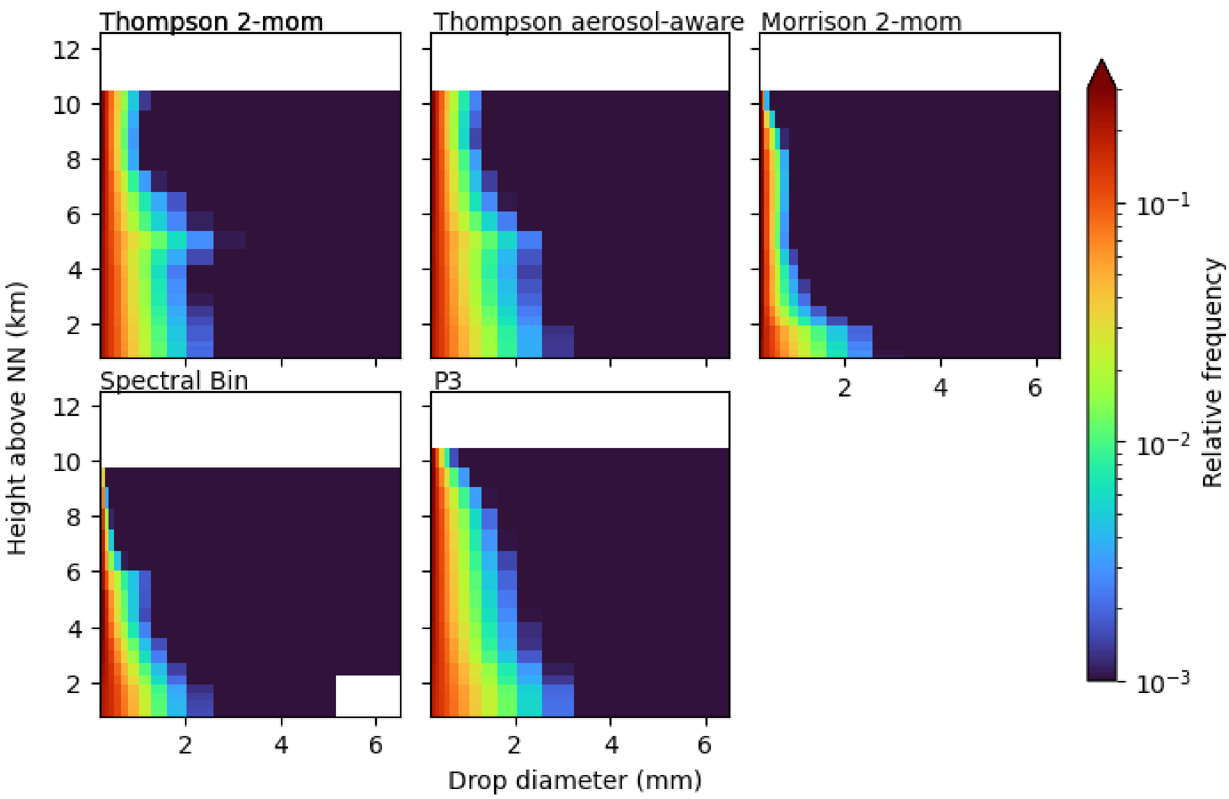
Section 4:

> *Using our framework, there are some challenges for the evaluation of the microphysics schemes performance. Using a large data set provides the possibility of a statistical evaluation. Thus, it can provide correct general overview of the schemes performance. On the other hand, considering long periods of time, multiple different weather situations produce convective cells of varying types. In our analysis, these are all analyzed together. This introduces ambiguities and some individual microphysical aspects might be smeared out. A solution would be a separation of different convective cloud types, e.g. by classifying into shallow, congestus or deep convective clouds using our 32 dBZ echo top height (e.g., Matsui et al., 2009). Furthermore, classifications into weak/strong forcing situations could be of interest, to analyze the effect of, e.g., frontal systems on the distribution of radar signals. This will be addressed in a future application of this framework.*

3) rain drop-size distributions

Probably the most robust finding in this study is the variability of rain-DSD related radar signals (ZDR and DWR) among different microphysics schemes as seen in Figures 5, 6, and 7. Above-melting-zone evaluations tend to have more uncertainties in the forward model (described in next). To augment your finding in radar signals and discussion, it's much better to directly examine simulated rain drop size distribution profiles (like CFAD format) from different microphysics schemes. This should not be a difficult task. (it's much better if you have disdrometer observations!)

We provided a CFAD of rain drop size distributions for convective cells (Appendix B):



The following passage was added to section 3.3:

> *In order to separate the analysis into reasons due to differences in the underlying modeled microphysics and due to different processing in the forward simulator, we examined rain particle size distributions directly produced by the NWP model (Rain PSD CFAD in Appendix B.) The FSBM scheme provides the drop size distributions over a number of size bins, for the other schemes we calculated the distributions according to the schemes parameterization. Only model grid boxes that were flagged as a convective cell by the TINT cell tracking are considered. The rain PSD CFAD confirms the findings of the ZDR CFAD: the two Thompson schemes simulate large rain drops from the surface up to the melting layer height and even above, while the Morrison scheme produces large rain drops only at the surface and the FSBM produce the highest frequency of small drops.*

4) uncertainties of forward model

Details and uncertainties in assumptions of the forward model (CR-SIM) are not discussed. In order to represent simulated microphysics in polarimetric observables, one must assume particle shape and orientation simultaneously in the forward model, because these are not "explicitly" simulated in most of the microphysics schemes in WRF. Following paper discusses and tests different assumptions. Please describe what kind of shape/orientation assumptions are made for each microphysics in Section 2.4 (or Appendix), and related discussions in Section 3.3.

Matsui, T., Dolan, B., Rutledge, S. A., Tao, W.â🚹🛢K., Iguchi, T., Barnum, J., & Lang, S. E. (2019). POLARRIS: A POLArimetric Radar Retrieval and Instrument Simulator. *Journal of Geophysical Research: Atmospheres*, 124. https://doi.org/10.1029/2018JD028317

We have not been clear on the assumptions of the radar forward model (CR-SIM). The section of CR-SIM (2.4) has been appended with the assumptions that CR-SIM is making regarding particle shapes, particle orientation and dielectric constants for each microphysics schemes:

The dielectric constant of water is 0.92. Solid phase hydrometeors are assumed to be dielectric dry oblate spheriods and are represented as a mixture of air and solid ice. The refractive index hence depends on the hydrometeor density and is computed using the Maxwell-Garnet (1904) mixing formula. There are no mixed phased particles simulated. This means mixed phase radar signatures (for example the "bright band", Austin and Bemis, 1950) will not be reproduced by the simulation. In order to simulate polarimetric radar observables, a radar forward simulator must assume particle shapes and particle orientation. The particle orientation assumptions are the same for all schemes. It is assumed that the particle orientations are 2D Gaussian distributed with zero mean canting angle as in Ryzhkov et al. (2011). The width of the angle distributions is specified for each hydrometeor class: 10° for cloud, rain, and ice and 40° for snow, unrimed ice, partially rimed ice, and graupel. Regarding the shape assumptions, cloud droplets are simulated as spherical (aspect ratio of 1) and raindrops are simulated as oblate spheriods with a changing axis ratio dependent on the drop size according to Brandes et al. (2002) in all schemes. For ice hydrometeor classes, the same aspect ratio assumptions are applied for all schemes except the P3 scheme: cloud ice is assumed as oblate with a fixed aspect ratio of 0.2. Snow is assumed as oblate with a fixed aspect ratio of 0.6. Graupel is assumed to be oblate with an aspect ratio that is changing from 0.8 to 1, dependent on the diameter and according to Ryzhkov et al. (2011):

$$ar = 1.0 - 0.02 \qquad \text{if } D < 10\ \text{mm}\,,$$
$$ar = 0.8 \qquad \text{if } D > 10\ \text{mm}\,.$$

The P3 scheme does not provide the standard ice hydrometeor classes. Instead, the aspect ratio of small ice (spherical, fixed aspect ratio of 1), unrimed ice (oblate, fixed aspect ratio of 0.6), partially rimed ice (oblate, fixed aspect ratio of 0.6) and graupel (spherical, fixed aspect ratio of 1) is assumed by CR-SIM. This means in comparison to the other schemes that the P3 simulation deviates for small ice (aspect ratio of 1 in P3, while cloud ice in other schemes is assumed to have an aspect ratio of 0.2) and graupel (0.8 - 1 in other schemes, while graupel particles in P3 are assumed to have an aspect ratio of 1). Resulting differences in the radar signal are discussed in the result section 3 whenever it might influence the simulated radar signal.

We further added a small abstract in the conclusions acknowledging uncertainties produced by the radar forward operator:

Furthermore, there are uncertainties connected to the radar forward simulator applied. To calculate scattering characteristics, assumptions have to be made including the particles aspect ratio, orientations, shape and more. The variability of the simulated signals is reduced by applying fixed relations compared to the potential variability of shapes, *orientations and aspect ratios in nature. In addition, the radar forward simulator applied in our study does not consider mixed phase particles. This means that, e.g., effects such as the bright band where particles melt cannot be reproduced by the simulations. To circumvent some ambiguities introduced this way, the comparison could be extended from radar signal space to cloud hydrometeor space. I.e., retrieved hydrometeor classes can be compared to simulated ones.*

Finally, we added multiple parts in the discussion and conclusions that relate forward simulator details to the results:

Section 3.2

Given that the forward simulator applied in this study does not consider wet particles, we find the high bias in Z exists even without considering wet graupel and comes mostly from rain, suggesting PSDs that contain too many large rain drops compared to the observations.

Section 3.3:

Furthermore, 3) the observed variability of ZDR is possibly not correctly captured by the radar forward simulator which has to assume fixed distributions of particle orientations as well as a fixed aspect ratio of the particles.

Section 3.3:

All schemes assuming spherical cloud ice or with other dominating spherical hydrometeor classes at these heights show small ZDR. This is true for the P3 small ice fraction for which the forward simulator assumes spherical aspect ratio of 1. In the Thompson schemes, the assumed aspect ratio by the forward simulator is 0.2, suggesting that other hydrometeor classes with lower ZDR like snow or graupel dominate the signal. Only for FSBM and Morrison (aspect ratio 0.2) cloud ice dominates the signal. The stronger signal in FSBM and Morrison is not a result of different density assumptions, because both, the FSBM and Morrison scheme assume lower density of cloud ice compared to Thompson. The observations do not show increased ZDR at these heights. This could either mean that 1) there are no large cloud ice particles observed, 2) that the signal is dominated by other more spherical particles in the observations, or 3) that the assumed aspect ratio of 0.2 by the radar forward operator is unrealistic and the observed particles are more spherical in nature.

Section 4:

This could either be a result of simulated cloud ice particles being too large or too many, but this could also be a result of the assumed flat cloud ice shape with an aspect ratio of 0.2.

Minor Comments/Suggestions

Line 25: “the huge number” -> “a large number”

Changed as suggested.

Line 26: “on scales of μm to mm and” -> “on scales of mm or smaller” for consistency. In fact, microphysics processes occur less than the scale of micron, such as ice crystallization processes.

Changed as suggested.

Lin 82: “with a sound statistical basis” ?? I don’t understand.

By sound statistical basis we mean a large sample size. We changed the phrase for clarification:

> *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 large sample size.*

Line 89: “separate the microphysical impacts from possible feedbacks.” I agree. But more bottom line, I would argue whether your set of numerical weather model resolved dynamics or not with 2km horizontal grid spacinig.

There is a misunderstanding here. The middle domain of our model setup has a 2 km horizontal grid spacing. The inner domain that we used for the analysis has a grid spacing of 400 m. We assume you refer to the cloud dynamics which we believe are resolved at a grid spacing of 400 m. That’s why we left this sentence as it is. Perhaps the comparison to current operational weather models (line 150) led to the confusion of our horizontal grid spacing. We added a sentence clarifying that our grid spacing is better than current operational numerical weather prediction models and is rather representing the future generation of NWP models in section 2.2:

> *Currently, operational limited area weather models operate at 2 km grid spacing (e.g., 2.8 km in COSMO-DE of the German Weather Service; Baldauf et al., 2011) which means our inner domain has a resolution that is effectively about 5 times higher and should be representing the future generation of operational limited area weather models.*

Line 108: “frequency” -> “frequencies”

Changed as suggested.

Line 149 and repeat the same: “a horizontal resolution of 10 km,” must be replaced by “horizontal grid spacing of 10km”. Numerical atmospheric model does not have 10km resolution with 10km of horizontal grid spacing. Effective dynamic resolutions are x5 ~ x10 of horizontal grid spacing in numerical dynamic core. Apply this correction elsewhere in the manuscript.

Pielke, R. A. (1991). A Recommended Specific Definition of “Resolution”, Bulletin of the American Meteorological Society, 72(12), 1914-1914. Retrieved Oct 30, 2021, from https://journals.ametsoc.org/view/journals/bams/72/12/1520-0477-72_12_1914.xml

Thank you for this input. Even though we think that ‘resolution’ is a widely accepted term instead of grid spacing, we understand the logic behind your comment and adapted the definition of the reference throughout the manuscript concerning the model grid descriptions.

Line 159: Please briefly describe other physics options, such as land surface, PBL, and radiation schemes.

The namelist of WRF is added as a supplement where all options used can be seen. The requested information about land surface, PBL and radiation scheme has been added to section 2.2:

> *Other physics options include the Noah Land Surface model (Ek et al., 2003; Chen and Dudhia, 2001), the MYNN2 planetary boundary layer scheme (Mellor-Yamada scheme by Nakanishi and Niino; Nakanishi and Niino, 2006) and the RRTMG radiation scheme (rapid radiative transfer model for general circulation models; Iacono et al., 2008). For any other options, please refer to the WRF namelist that is provided as a supplement to this manuscript.*

Line 203: Did you store and use all 33bin of hydrometeor classes to calculate radar observables in CR-SIM?

Yes, we stored and used all bins of all hydrometeor classes to calculate the radar observables for the fast spectral bin simulations. We also stored the aerosol bins (43), but these are not used by CR-SIM. However, the spectral bin scheme uses shared bins for rain / cloud droplets (First 17 bins for cloud droplets, second 16 bins for rain) and cloud ice / snow (First 17 bins for cloud ice, second 16 bins for snow). The output for graupel consists of the full 33 bins. The data is saved at our institute and available on request. We don't think this information is relevant for the reader which is why we did not change the phrasing at this point.

Line 286: "but none of them as pronounced as in the observations." Well, this is typical situations that relatively coarse-resolution model won't be able to resolve tiny cells. So you are running with 2km horizontal grid, meaning that you can resolve convective features in 10km or 20km well, but never be able to resolve 2km-size of convection, which tend to have shallower echo-top heights. So, don't blame to microphysics, but model dynamic core and grid spacing you chose.

This is again connected to a misunderstanding. We use a 400 m horizontal grid spacing. That means we are able to resolve convective cells at 2 km or 4 km in size. However, the point still stands: it is likely that we miss the very small convective cells anyways which correlate to the lower echo-top heights. It was not our intention to blame the cloud microphysics for this, as this is a feature in all simulations independent of the cloud microphysics. We slightly rephrased the sentence to clarify the meaning:

> *All NWP simulations independent of the microphysics scheme are able to reproduce a peak at a similar altitude but none of them as pronounced as in the observations.*

Furthermore, we added another sentence in the following abstract to emphasize that we don't blame the microphysics for this effect:

> *This is independent of the chosen cloud microphysics scheme and mainly a result of the missing small-scale cells in the simulations which is indicative of a resolution effect: the*

very small cell heights correspond to tiny cells that we might not be able to resolve even with our 400 m grid spacing.

Line 313: “Contoured frequency by altitude distributions” -> “Contoured frequency of altitude diagram”

Changed as suggested, also applied elsewhere in the manuscript.

Line 322: “image 5” -> “Figure 5”?

Changed as suggested.

References

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, 1950.

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.

Brandes, E. A., Zhang, G., and Vivekanandan, J.: Experiments in rainfall estimation with a polarimetric radar in a subtropical environment, Journal of Applied Meteorology, 41, 674–685, 2002.

Chen, F. and Dudhia, J.: Coupling an advanced land surface–hydrology model with the Penn State–NCAR MM5 modeling system. Part I: Model implementation and sensitivity, Monthly weather review, 129, 569–585, 2001.

Ek, M., Mitchell, K., Lin, Y., Rogers, E., Grunmann, P., Koren, V., Gayno, G., and Tarpley, J.: Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta model, Journal of Geophysical Research: Atmospheres, 108, 2003

Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., and Collins, W. D.: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models, Journal of Geophysical Research: Atmospheres, 113, 2008.

Jung, Y., Xue, M., and Zhang, G.: Simulations of polarimetric radar signatures of a supercell storm using a two-moment bulk microphysics scheme, Journal of Applied Meteorology and Climatology, 49, 146–163, 2010.

Kumjian, M. R.: The impact of precipitation physical processes on the polarimetric radar variables, The University of Oklahoma, 2012.

Maxwell-Garnet, J. C.: Colours in metal glasses and in metallic films, Phil. Trans. R. Soc. Lond, A, 203, 385–420, 1904.

Matsui, T., Zeng, X., Tao, W.-K., Masunaga, H., Olson, W. S., and Lang, S.: Evaluation of long-term cloud-resolving model simulations using satellite radiance observations and multifrequency satellite simulators, Journal of Atmospheric and Oceanic Technology, 26, 1261–1274, 2009.

Nakanishi, M. and Niino, H.: An improved Mellor–Yamada level-3 model: Its numerical stability and application to a regional prediction of advection fog, Boundary-Layer Meteorology, 119, 397–407, 2006.

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, 2017.

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, 2011.

Snyder, J. C., Bluestein, H. B., Dawson II, D. T., and Jung, Y.: Simulations of polarimetric, X-band radar signatures in supercells. Part I: Description of experiment and simulated ρhv rings, Journal of Applied Meteorology and Climatology, 56, 1977–1999, 2017.