5B.5 Intercomparison of simulations using 4 WRF microphysical schemes with dual-Polarization data for a German squall line

William A. Gallus, Jr.[1]
Monika Pfeifer[2]

[1]*Iowa State University*
[2]*DLR Institute of Atmospheric Physics*

1. Introduction

Improvements in computational power in recent years have allowed for increasing use of fully explicit numerical models for weather prediction. It is well-known that the use of convective parameterizations in models introduces substantial error, and the use of near-cloud resolving grid spacings may reduce the error since convective schemes can be eliminated. However, microphysical schemes may also introduce great uncertainty into forecasts (e.g., Jankov et al. 2005), and far fewer studies have been done to verify the results from different microphysical packages.

In this study, the ability of a fine grid version of the Weather Research and Forecasting (WRF) model to reproduce accurately the microphysical structure of a squall line is evaluated by comparing WRF simulations using four different microphysical schemes to detailed observations gathered by the DLR polarimetric radar POLDIRAD (Schroth et al., 1988) located near Munich, Germany. A squall line event which occurred on August 12, 2004 is simulated with each version of the model. Synthetic polarimetric radar scans are derived from the model forecasts employing the polarimetric radar forward operator SynPolRad (Pfeifer et al., 2004). The evaluation of the microphysical parameterization schemes is carried out comparing PPI and RHI scans of reflectivity and the spatial distribution of hydrometeor types. The hydrometeor types are derived applying a hydrometeor classification scheme (Hoeller et al., 1994) to the observed and simulated polarimetric radar quantities. Furthermore, the Ebert-McBride contiguous rain area method of verification is tested in a variety of ways on the reflectivity output from the simulations.

2. Data and Methodology

On August 12th, 2004 a cold front crossed Germany resulting in the development of severe thunderstorms with high reflectivities in southern Bavaria. The front developed into a squall line with a sharply defined convective line producing hail and a trailing region of stratiform precipitation. The squall line examined formed prior to 15 UTC and reached its maximum intensity around 17 UTC (Fig. 1), when it was oriented north-south and passing very near the radar. In the next 1-2 hours, the system began bowing as it moved eastward. Also during this time, a trailing stratiform rain region developed.

Corresponding author address: William A. Gallus, Jr., ISU, 3010 Agronomy Hall, Ames, IA. Email: wgallus@iastate.edu.

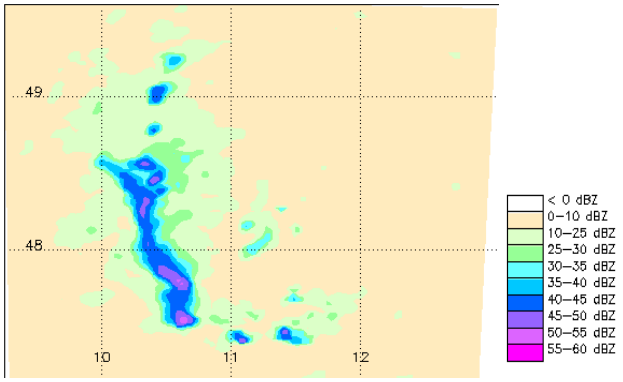

Figure 1: 1° PPI scan of reflectivity [dBZ] observed by POLDIRAD at 1659 UTC. The data is interpolated onto the WRF model grid for evaluation.

The WRF model version 2.1.2 with the ARW (Advanced Research WRF) dynamic core was run over a roughly 280 by 280 km inner domain centered near Munich, Germany, having 2.8 km grid spacing, nested within a larger domain of approximately 1000 by 1000 km having 8.4 km grid spacing. Simulations were integrated for 24 hours using 00 UTC 12 August GFS output for initialization and lateral boundary conditions. Thus, spin-up problems should be minimized since the primary squall line did not occur until after 12 hours into the simulations. No convective parameterization was used on either grid. The four microphysical schemes used included Lin et al (Lin et al. 1983, Chen and Sun 2002), older Thompson (Thompson et al. 2004), WSM6 (Hong et al. 2004), and WSM5 (Hong et al. 2004). The WRF runs will be referred to as Lin, WSM6, WSM5 and Thompson hereafter. A newer version of Thompson became available recently and preliminary results (not shown here) suggest some improvements over the older version. All versions of the model used the Dudhia shortwave and RRTM longwave radiation schemes, NOAH land surface model, and YSU PBL scheme.

All verification was done on the inner 2.8 km grid spacing domain. The observations were interpolated onto the model grid to allow comparisons at the same horizontal resolution. Subjective verification was performed using observed and synthetic PPI and RHI plots of reflectivity and a hydrometeor classification. In addition, objective verification was performed using a version of the Ebert-McBride technique (EMT) applied to MCSs in the central United States (Grams et al. 2006). The EMT identifies contiguous

rainfall areas (CRAs) which can be used to determine displacement errors within models. The EMT was applied at 18 UTC assuming no temporal error in the simulations, and also applied accounting for temporal errors.

3. Results

When the EMT was applied at 18 UTC to each simulation (Table 1), it found that only WMS6 did not have a high bias in the number of points with positive reflectivity values. Lin and WSM5 both were the worst with a roughly 60% overestimate of the number of points. Average observed reflectivities in the domain were around 5.6 dBZ, but were 9.0 in WSM6, 11.6 in Thompson, 13.2 in Lin, and 14.0 in WSM5. These results suggest all of the schemes produce too much reflectivity. However, the maximum values at any grid point at 18 UTC were relatively close to those observed. The observed maximum at this time was 48.7 dBZ, with all schemes having a minor overestimation, the worst being Thompson with 53.7.

	Obs	Lin	WSM6	WSM5	Thomp
points	3758	5935	3727	5758	4477
ave	5.61	13.15	9.03	13.96	11.55
max	48.7	51.0	49.1	49.9	53.7
corr		.031	.392	-.006	-.209
ETS		.139	.188	.090	-.040

Table 1: Total number of points with reflectivity above 0, average and maximum reflectivity (dBZ), correlation coefficient and ETS for the 4 model runs at 18 UTC.

Correlation coefficients were near zero or negative for all schemes but the WSM6, which did suprisingly well at 18 UTC, having a value of .392. The best Equitable Threat Score (ETS) for all reflectivity above zero in the entire domain was for WSM6. Thompson scored the worst.

Using 25 dBZ as the threshold for which the EMT scheme determines its CRAs (Table 2), Lin best depicted the squall line, with an overestimate of integrated reflectivity of around 20%. By contrast, the overestimates exceeded 90% for WSM6 and 250% for WSM5 and Thompson. These results suggest far too much reflectivity above 25 dBZ was present in those schemes. Displacement error was smallest, however, in WSM6, around 30 km to the NNW. For Lin, it was 55 km to the ENE, WSM5 was 65 km to the NE, and for Thompson, 80 km to the NNW. Thus in all runs, there was a tendency at this time to show too much reflectivity north of the main squall line. The worst RMS errors prior to accounting for displacement error were in Thompson, consistent with its high bias for reflectivity. The best RMS errors were in WSM6. After accounting for displacement error, Lin had the least RMS error, while Thompson still had the most. Lin also showed a dramatic shift in the correlation coefficient as displacement was taken into account. It had the worst correlation prior to shifting the CRA, and the best afterwards, a value of .668. This result suggests that the forecasted reflectivity structure was very good but a displacement error was present. An error decomposition provided by the EMT agrees with this analysis. It suggests the worst displacement error was present in Lin. The

	Obs	Lin	WSM6	WSM5	Thomp
Int-dBZ	3.77	4.42	7.36	13.76	14.35
Disp-E		.45	-.08	.42	-.19
Disp-N		.20	.25	.50	.76
RMS-orig		26.4	20.7	26.4	27.4
RMS-shift		12.4	16.2	17.3	22.2
Corr-orig		-.604	-.032	-.482	-.435
Corr-shift		.668	.410	.538	.143
Err-Disp		544.3	165.2	397.7	255.4
Err-Vol		16.1	34.6	105.4	100.9
Err-Patt		137.2	226.5	193.7	392.5

Table 2: CRA analysis results at 18 UTC for all 4 model runs. Int-dBZ is the integrated reflectivity ($x10^3$ dBZ km^2), Disp-E and –N are displacements in degrees of latitude (N) or longitude (E), RMS and Corr are the root mean square error and correlation coefficients before and after shifting the CRA to account for displacement error. Error decompositions are shown for displacement (Disp), volume (Vol) and pattern (Patt).

smallest displacement error was in WSM6, which also had the smallest total error. Total errors were much larger and relatively similar in the other 3 runs. Pattern errors were responsible for most of the error in WSM6 and Thompson, while displacement errors dominated Lin and WSM5. Overall, the 18 UTC verification generally suggests WSM6 performed the best at this time. It would be difficult to say which of the other three performed the worst.

Because timing errors can be responsible for problems as well in the forecasts, the EMT was also applied to the simulations allowing for a small shift in the timing of key events in the evolution of the squall line. In particular, the verification was performed at the time when the system was intense near Munich, and at a time when the system had a well-developed stratiform region. Regarding the comparison of peak convective organization, only WSM6 seemed to have the correct timing for when the squall line was intense and oriented north-south around 17 UTC. The other three schemes produced squall lines that developed too quickly, passing the Munich area around 16 UTC instead of 17 UTC. Figure 2 shows the reflectivities in the four model runs at the times their depictions best matched the 1659 UTC observations (Fig. 1). Nonetheless, considering this was 16 hours into the forecast, the one hour error seems relatively minor.

For the entire domain at the time of best convective line structure (Table 3), all 4 model versions underestimated the number of points with reflectivity. The average reflectivities in the domain were much closer to the observed value than in the 18 UTC analysis, with WSM5 and WSM6 having small overestimates (by less than 10%) and Lin and Thompson having slightly larger underestimates (12-18%). Maximum reflectivity observed was 52.0 dBZ, and all 4 model versions underestimated this, with values between 48 and 50 dBZ. Correlation coefficients were between .37 (Thompson) and .56 (WSM5). ETS values were highest for WSM6 and WSM5, with Lin only slightly lower. At this time, it would be difficult to state which scheme worked best, but the Thompson clearly performed the worst.

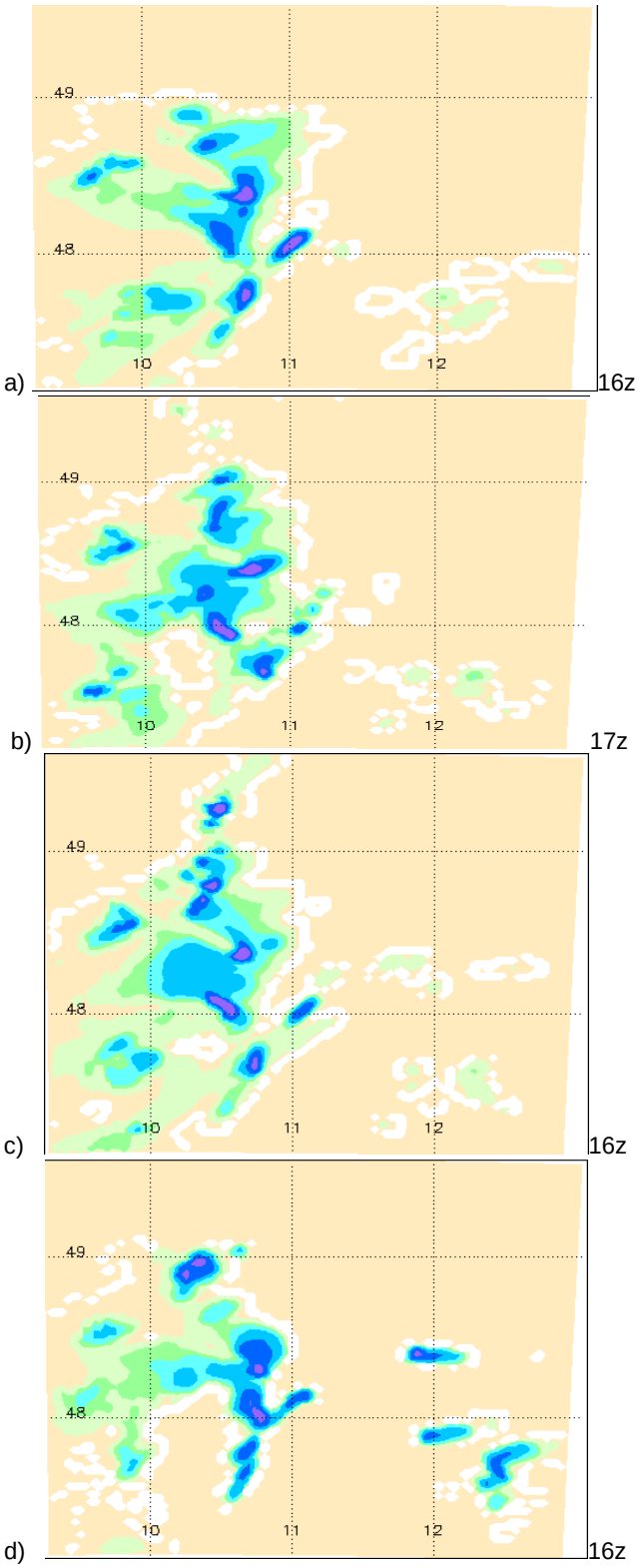

Figure 2: Simulated reflectivity for a) Lin, b) WSM6, c) WSM5, and d) Thompson, using scale shown in Fig. 1.

Regarding CRAs at that time, all models overestimated the areal coverage of reflectivity above 25 dBZ. Considering that all models underestimated the areal coverage of any positive reflectivity within the full domain, this implies the models are aggressive at producing relatively high reflectivity values, but lacking in lighter reflectivity areas at this time. Displacement errors were much smaller in all four runs, which makes sense since the temporal error was subjectively accounted for by choosing times when the model output best resembled the observations. In all cases the displacements were less than 27 km. Changes in RMS error and correlation coefficient after accounting for the displacement error were much smaller than at 18 UTC for a similar reason. The best RMS errors were present in WSM6 and WSM5. The best correlation coefficients were in WSM5. In all 4 cases, pattern errors were by far the biggest contributor to the total error, and the total error was smallest in WSM5 and largest in Thompson. A sensitivity test was performed for the model data at these times to determine what role the reflectivity threshold parameter played in the EMT results. In a test using 30 dBZ instead of 25 dBZ, noticeable changes occurred in some parameters for some model runs but the changes were not substantial enough to change the results described above.

A stratiform region was well-developed in the radar data by 19 and 20 UTC (Fig. 3). In the model simulations, its evolution was dependent on the microphysical scheme used (Fig. 4). Subjectively, the best agreement with observations in the general location of the stratiform rain took place at 20 UTC in Thompson and WSM5, and at 19 UTC in Lin. In WSM6, the best agreement was between the 20 UTC model output and 19 UTC observations, implying a lag in development in the model. None of the model runs showed as much organization to the stratiform region as observations indicated. For the entire domain at the time of best depiction of a stratiform region, the areal coverage of positive reflectivity evidenced very different trends compared to the areal coverage at time of maximum intensity of the convective line. All 4 model versions had a low bias in the total number of points with positive reflectivity, with the smallest underestimate of around 13% in Lin, and the worst being almost 50% in Thompson. Average reflectivities, however, were too high in all runs except Thompson, but even in Thompson, the underestimate of average reflectivity, around 30%, was less than the underestimate of areal coverage. These trends suggest that all of these microphysical schemes fail to produce large enough areas of low reflectivity, and continue to show too many areas of

	Obs	Lin	WSM6	WSM5	Thomp
points	3437	2388	2620	2811	1899
ave	5.90	5.23	6.15	6.22	4.95
max	52.0	48.7	48.7	49.9	47.7
corr		.489	.530	.556	.368
ETS		.287	.314	.312	.168

Table 3: Total number of points with reflectivity above 0, average and maximum reflectivity (dBZ), correlation coefficient and ETS for the 4 model runs at the time of best subjective agreement of forecast to 17z observations.

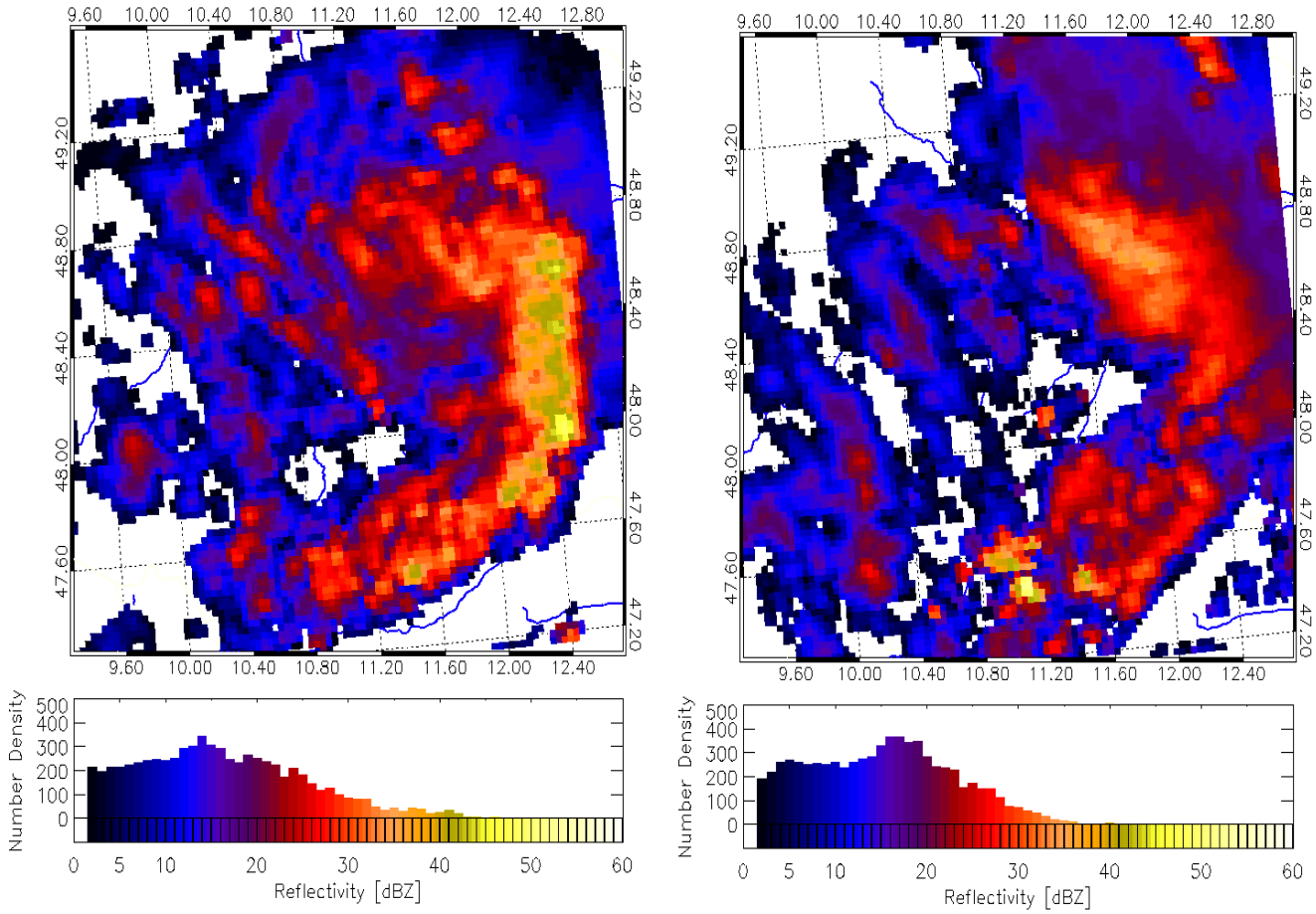

Figure 3: Reflectivity values and histograms of occurrence frequency from Munich radar at 1906 UTC (top) and 2007 UTC (bottom).

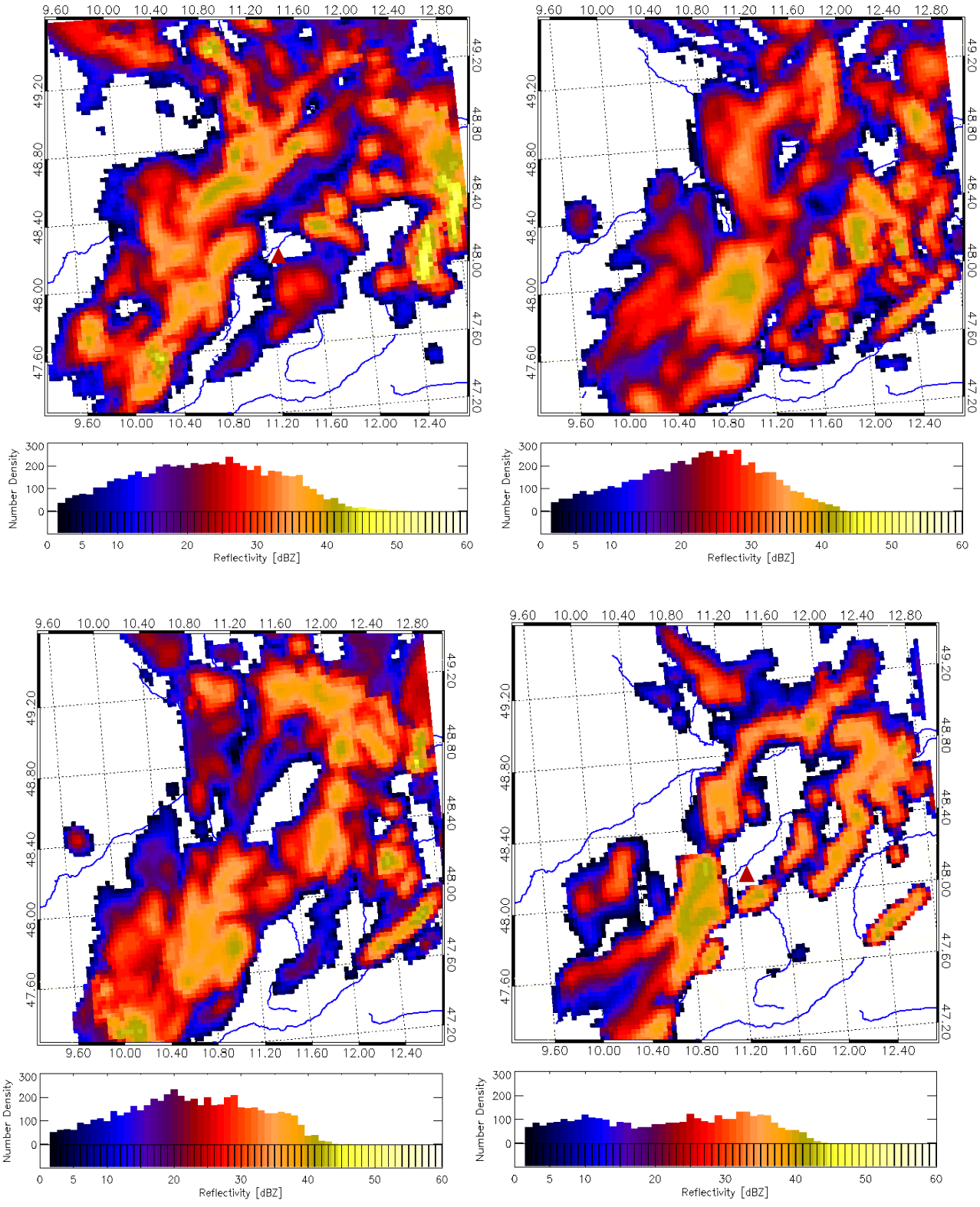

Figure 4: Simulated reflectivity for Lin at 19 UTC (upper left), and WSM6 (upper right), WSM5 (lower left) and Thompson (lower right) at 20 UTC, along with histogram of frequency of occurrence of reflectivity values.

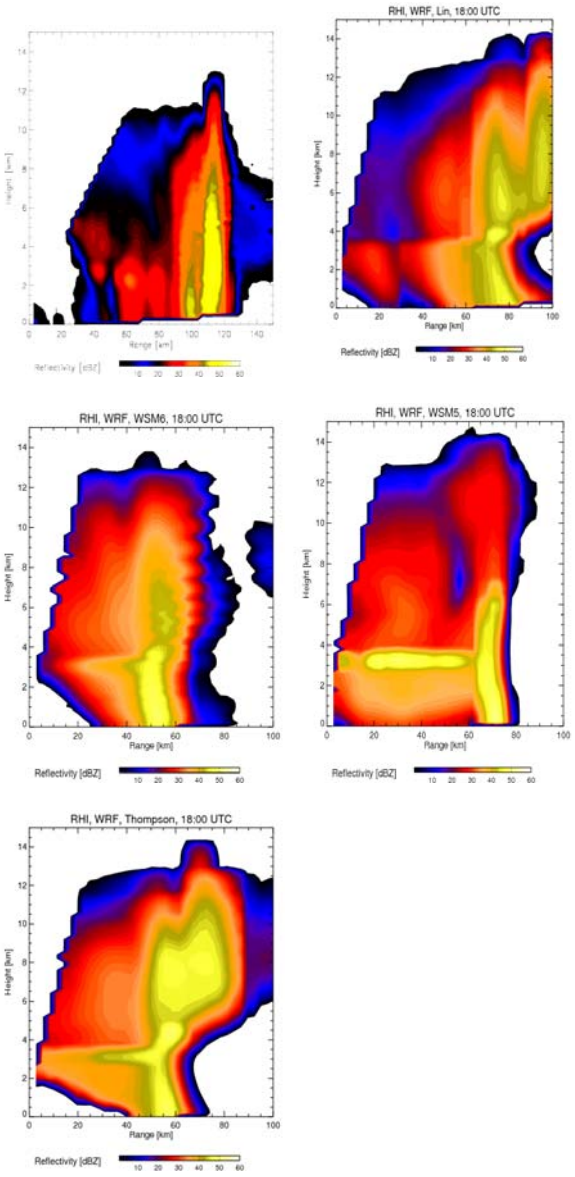

Figure 5: RHI scans of observations (upper left), Lin (upper right), WSM6 (middle left), WSM5 (middle right) and Thompson (bottom).

moderately high reflectivity. A comparison of histograms in Figs. 3 and 4 show that the models depict too little coverage of reflectivities below roughly 20 dBZ and too much coverage above. Peak reflectivities, however, are less than the observed value in all runs except Lin. The best correlation coefficient, .438, occurs in WSM6. Other correlation coefficients are much smaller, ranging from .258 in Lin to .124 in Thompson. ETSs behave similarly. For the full domain at these times, Thompson seemed to have the worst scores in general. No one scheme stood out as best.

For the CRAs identified at this time, all runs continued to overestimate the area of reflectivity above 25 dBZ. Northward displacements exist in all 4 runs, with significant eastward displacements in all but the WSM5. The eastward displacements might suggest that the model versions failed to develop enough stratiform rainfall to the rear of the system. The best displacement error was around 20 km to the NNE in Thompson. The worst was around 45 km to the NE in Lin. The smallest RMS errors occurred in WSM5, around 11.65 after accounting for displacement error. The worst error was present in Lin, around 17.31. All correlation coefficients were negative prior to shifting the CRA; after the shift, the WSM5 value improved to .367 while the other runs had more modest improvements to around 0. Pattern errors dominated the error decomposition in all runs, but were relatively less severe in WSM5.

Errors were also apparent in the RHI (range-height indicator) scans of the simulated reflectivities (Fig. 5). These images were created around 18 UTC by taking vertical cross-sections through the systems at a point where the convective line subjectively appeared to be most intense. All runs seemed to have too much water mass at high levels in the stratiform regions, although the problem was less severe with Lin. Lin and Thompson had too much high reflectivity at upper levels ahead of the main convective line. All schemes except Lin also had too much reflectivity at low levels in the stratiform region. The bright band behind the line was too intense in all cases except Lin, where it was a little too weak. The too intense bright band is especially noticeable in WSM5. Within the convective line region, Lin was too weak with reflectivity in both the rain and ice regions, WSM6 was too weak aloft but acceptable in the rain region, Thompson was too intense aloft and WSM5 agreed best with observations. Both Lin and Thompson incorrectly depicted a bright band in this region. Overall, Lin seemed to best capture the structure of the reflectivity, despite having an incorrect depiction of a convective cell ahead of the line at high levels. It was the only run that came close to showing the reflectivity structure correctly behind the convective line.

Dual-polarization radar can be used to diagnose the hydrometeor species present in the data, and therefore, this radar offers one of the few ways available to verify qualitatively the mixing ratio contents of different species produced by microphysical schemes. Figure 6 shows the hydrometeor species dominating in the observations and each of the 4 model runs. Lin, WSM6, and Thompson show the same general trends, but in all 3 cases, far more graupel is produced than is present in the radar data. The only run without this problem is WSM5, which does not include a graupel classification. It should be noted that the T-matrix scattering code used by SynPolRad simulates attenuation, and as with any radar, incorrect hydrometeor classifications may result in bright band regions or areas with strong attenuation, such as the area just behind the convective line. Anomalous depiction of wet and dry hail in Fig. 6 near the stratiform bright band area and the heavy rain of the convective line can be seen. None of the model runs depicts the heavy rain that the radar data shows just ahead of the main convective core, instead showing all heavy rain to lie behind it.

4. Summary and Conclusions

Four high resolution WRF simulations using different microphysical schemes were compared with detailed observations of a German squall line gathered by the DLR polarimetric radar POLDIRAD. Synthetic polarimetric radar scans were derived from the model forecasts employing the polarimetric radar forward operator SynPolRad. The evaluation of the microphysical parameterization schemes was carried out comparing PPI and RHI scans of reflectivity and the spatial distribution of hydrometeor types.

All microphysical schemes were found to share some problems. For instance, all resulted in runs that had too little areal coverage of reflectivity early in the development of the convective line, but quickly produced too intense reflectivities on average throughout the domain. At all times evaluated, all runs seemed to produce too many areas with intense reflectivity (generally over 25 dBZ) and not enough areas with light reflectivity. All runs overestimated reflectivity in the stratiform portion of the system as it matured. The hydrometeor classification scheme also showed that all runs that included a graupel classification incorrectly depicted graupel as the dominant hydrometeor aloft throughout the system when snow was the dominant observed hydrometeor.

Other deficiencies affected a subset of the microphysical schemes. Bright bands were incorrectly shown within the convective lines of the Lin and Thompson runs. Reflectivities within ice were too small in the convective lines of the Lin and WSM6 runs, but too large in the Thompson run. Reflectivities within the convective rain region were too small in the Lin run. All runs except Lin had too much reflectivity in the stratiform anvil. The bright band was overestimated in WSM5 and Thompson.

Future work should investigate the sensitivity of the results to the grid spacing used in the model, and extend the sample to other cases including other convective modes. If the results found for this squall line are robust, techniques like SynPolRad could be used to help model developers make changes in their microphysical schemes to better agree with radar observations. The impacts of such changes on rainfall, however, would have to be investigated.

5. Acknowledgments

This research was funded in part by National Science Foundation grant ATM-0537043.

6. References

Chen, S.-H., and W.-Y. Sun, 2002: A one dimensional, time dependent cloud model. *J. Meteor. Soc. Japan*, **60**, 99-118.

Grams, J. S., W. A. Gallus, Jr., L. S. Wharton, S. E. Koch, A. Loughe, and E. E. Ebert, 2006: The use of a modified Ebert-McBride technique to evaluate mesoscale model QPF as a function of convective system morphology during IHOP 2002. *Wea. Forecasting*, **21**, 288-306.

Hoeller, H. , V. Bringi, J. Hubbert, M. Hagen, P. F. Meischner, 1994: Life cycle and precipitation formation in a hybrid-type hailstorm revealed by polarimetric and Doppler radar measurements. J. Atmos. Sci. , **51**, 2500-2522.

Hong, S.-Y., J. Dudhia, S.-H. Chen, 2004: A revised approach to ice microphysical processes for the bulk parameterization of cloud and precipitation. *Mon. Wea. Rev*., **132**, 103-120.

Jankov, I., W. A. Gallus, Jr., B. Shaw, and S. E. Koch, 2005: On the impacts of different WRF physical parameterizations and their interactions on warm season MCS rainfall. *Wea. Forecasting*, **20**, 1048-1060.

Lin, Y.-L., R. D. Farley, and H. D. Orville, 1983: Bulk parameterization of the snow field in a cloud model. *J. Appl. Meteor*., **22**, 1065-1092.

Pfeifer, M., G. Craig, M. Hagen, and C. Keil , 2004: A polarimetric radar forward operator. *Proceedings of ERAD*, **2**, 494-498

Schroth. A. C. , M. S. Chandra, and P. F. Meischner, 1988: A C-band coherent polarimetric radar for propagation and cloud physics research. *Journal of Atmospheric and Oceanic Technology*, **5**, 804-822

Thompson, G., R. M. Rasmussen, and K. Manning, 2004: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part I: Description and sensitivity analysis. *Mon. Wea. Rev*., **132**, 519-542.

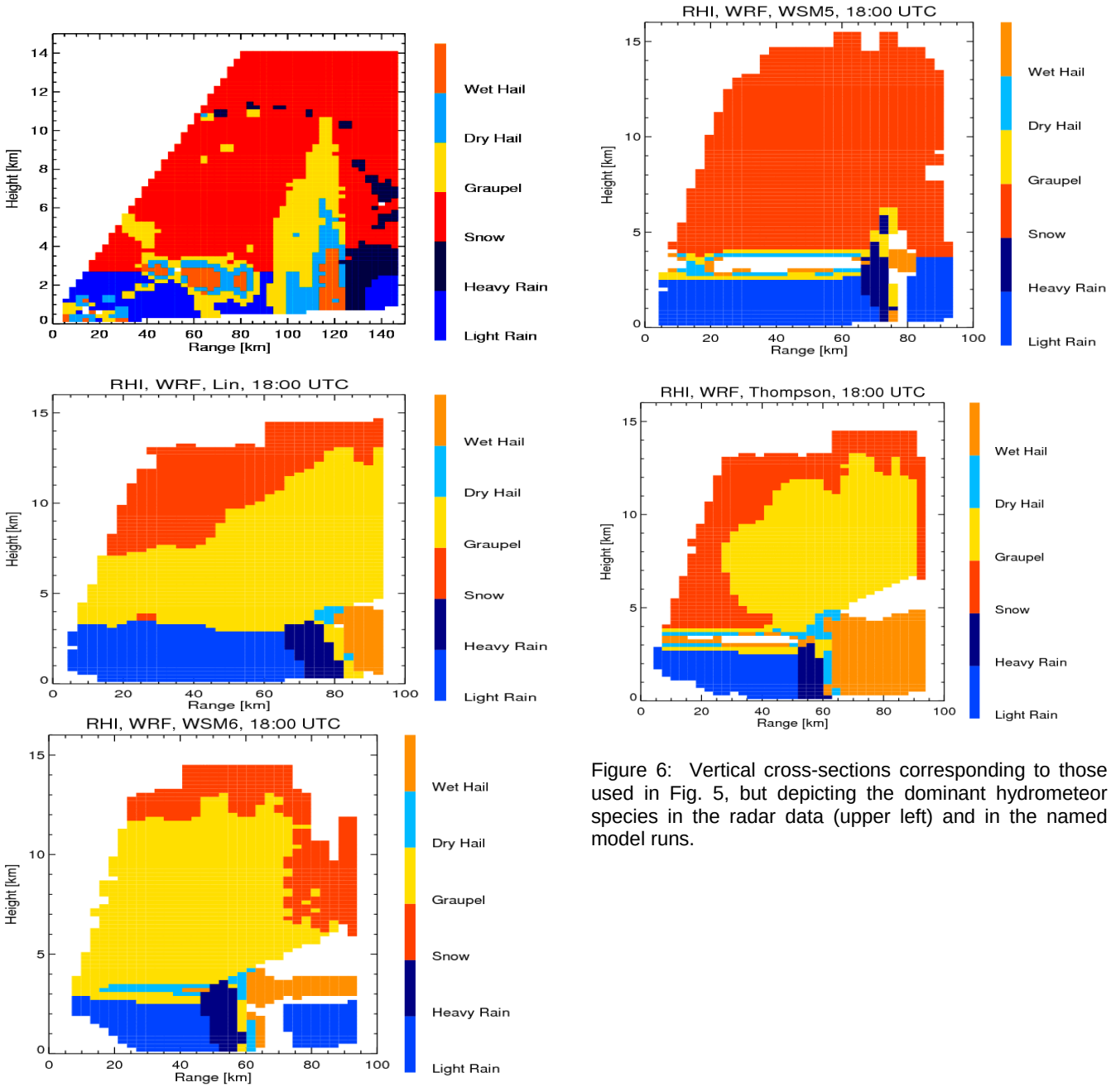

Figure 6: Vertical cross-sections corresponding to those used in Fig. 5, but depicting the dominant hydrometeor species in the radar data (upper left) and in the named model runs.