Using Radar Reflectivity to Evaluate the Vertical Structure of Forecast Convection

MARIUSZ STARZEC AND GRETCHEN L. MULLENDORE

Department of Atmospheric Sciences, University of North Dakota, Grand Forks, North Dakota

PAUL A. KUCERA

Research Applications Laboratory, National Center for Atmospheric Research,^a Boulder, Colorado

(Manuscript received 25 April 2018, in final form 15 October 2018)

ABSTRACT

Several months of regional convection-permitting forecasts using two microphysical schemes (WSM6 and Thompson) are evaluated to determine the accuracy of the simulated convective structure and convective depth and the impact of microphysical scheme on simulated convective properties and biases. Forecasts are evaluated by using concepts from object-based approaches to compare the three-dimensional simulated reflectivity field with the reflectivity field as observed by radar. Results from analysis of both schemes reveals that forecasts generally perform well near the surface but differ considerably aloft both from observations and from each other. Forecasts are found to contain too many convective cores that are individually larger than in the observations, with at least double the number of observed convective cores reaching the midtroposphere (i.e., 4–8 km). Although the number of cores is overpredicted, WSM6 cores typically contain lower simulated reflectivity values than the observations, and the regions of highest reflectivity values do not extend far enough vertically. Conversely, Thompson cores are found to have significantly higher reflectivity values within cores, with the strongest intensities extending higher than in the observations and having magnitudes higher than any observed cores. Forecast reflectivity distributions within convective cells are found to contain more spread than in the observations. The study also assessed the uncertainty in simulated reflectivity calculations by using a second commonly utilized method to calculate simulated reflectivity. The sensitivity analysis reveals that the primary conclusions with each method are similar but the variability generated by using different simulated reflectivity calculations can be as pronounced as when using different microphysical schemes.

1. Introduction

Model simulations are commonly used for both operational forecasting and to investigate processes in the atmosphere when inadequate observations exist or are hard to obtain, such as involving deep moist convection. In particular, the height, depth, and internal characteristics of deep moist convection are important storm factors for the operational and research communities. Storm height can be an indicator of convective strength, severity, and potential for lightning and hail generation (e.g., Held 1978; Ushio et al. 2001; Donavon and Jungbluth 2007; Pessi and Businger 2009; Yang and King 2010). The depth and internal structure also influence the generation

location and magnitude of convectively induced turbulence, which has an impact on aviation (e.g., Lane et al. 2003; Barber et al. 2018). The internal storm dynamics and microphysics are closely tied together and affect attributes such as precipitation intensity, latent heat release, and mass transport (e.g., Adler and Mack 1984; Houze 1989; Mullendore et al. 2005; Powell and Houze 2015). For example, modeling studies of convective mass transport rely on atmospheric chemical models to simulate tracer and chemical transport, reactions, and dilution. While atmospheric chemical models provide three-dimensional detail to transport processes, these models need to first simulate properly the internal dynamics of deep convection to accurately disperse such tracers. Knowledge of the height of the upper-level storm detrainment, as dictated by storm dynamics, is crucial as transported boundary layer air has different radiative and climatic effects depending on the altitude of the transported mass (e.g., Dickerson et al. 1987; Barth et al. 2007; Anderson et al. 2012). Unfortunately, convection-allowing

^aThe National Center for Atmospheric Research is sponsored by the National Science Foundation.

Corresponding author: Mariusz Starzec, mariusz.starzec@my.und.edu

simulations are largely unconstrained with regard to storm height and internal characteristics; therefore, modeling studies typically have to subjectively compare the simulated convection to its observed counterpart to see whether the simulated storm depicts similar storm features. Furthermore, as regional convection-permitting climate model simulations and down-scaling of simulations become utilized more, it is increasingly important to correctly simulate the three-dimensional storm structure to accurately balance the atmospheric energy budget. Objective methods are needed to evaluate the three-dimensional structure of simulated storms.

A form of evaluation that accounts for spatial differences is necessary in high-resolution forecasts as traditional skill scores have limited usefulness when spatial anomalies exist between forecasts and observations (e.g., [Mass et al. 2002](#)). It is known that the exact timing and location of convective initiation is very difficult to predict due to the chaotic nature of the boundary layer. Consequently, using traditional point-to-point verification often leads to a subjectively good forecast being deemed objectively bad (e.g., [Baldwin and Kain 2006](#); [Mittermaier 2014](#)). Therefore, to evaluate the representation of convection within convection-allowing simulations, the procedure should not be focused on precise location but should focus on evaluating convective characteristics, such as intensity, depth, and size. One methodology that does not emphasize precise locations and focuses more on such features is an object-based framework (e.g., [Ebert and McBride 2000](#); [Davis et al. 2006](#); [Ebert and Gallus 2009](#)). An advantage of object-based methods is that they focus on features within the analysis field rather than the entire field itself, meaning that regions of interest (i.e., objects) are defined based on the criteria specified by the user performing the analysis. The object-based method enables a comparison of similar objects and their attributes in both model simulations and observations.

At convection-allowing scales, object-based methods have been used to investigate various parameters such as updraft helicity ([Clark et al. 2012](#)), brightness temperature ([Griffin et al. 2017](#)), and convective initiation ([Burghardt et al. 2014](#)), but the predominant focus has been on the evaluation of precipitation fields (e.g., [Davis et al. 2009](#); [Ebert and Gallus 2009](#); [Gallus 2010](#); [Johnson and Wang 2013](#); [Johnson et al. 2013](#); [Clark et al. 2014](#); [Cai and Dumais 2015](#)). The general consensus from the aforementioned studies indicates that high-resolution model simulations commonly overforecast the number of precipitating objects, precipitating objects are frequently too large horizontally, and precipitation tends to have a high intensity bias. While knowledge of biases in the simulated precipitation fields provides important insights for operational forecasting, it is

also important to determine what processes may lead to those biases. Since the precipitation field is temporally averaged, potentially important features and details may be smoothed out. As a result of temporal smoothing, different forms of convection can produce a similar precipitation field while being dynamically and/or microphysically distinct from each other. In contrast, the reflectivity field is not temporally averaged and provides an instantaneous representation of convective processes enabling a more direct comparison of storm features.

Comparisons of the simulated and observed reflectivity fields have been successfully used to investigate several model characteristics in the past; however, studies have evaluated the composite reflectivity field or the reflectivity field at one level, limiting the usage of vertical information (e.g., [Kain et al. 2008](#); [Schwartz et al. 2009](#); [Jensen et al. 2010](#); [Kain et al. 2010](#)). Analyses that evaluated vertical convective structure have typically been limited to case studies or short periods of time (e.g., [Rogers et al. 2007](#); [Van Weverberg et al. 2011](#); [Caine et al. 2013](#); [Min et al. 2015](#)); nevertheless, these studies have revealed important model biases. For example, [Caine et al. \(2013\)](#) analyzed several days of model simulations and found that the simulations tended to have higher storm tops than the observations. Similarly, [Van Weverberg et al. \(2011\)](#) found simulated reflectivity values that were too high in their simulations of supercell and multicell cases.

In this study, the observed reflectivity field at multiple heights is utilized to evaluate several months of daily convection-allowing forecasts over the northern Great Plains. The purpose of this study is to demonstrate the usefulness of utilizing three-dimensional reflectivity for model evaluation, to investigate the accuracy of forecast convective depth and internal vertical convective structure, and to illustrate how convective biases change with height. Two different microphysical schemes are analyzed to investigate the variability generated in convective forecasts by switching schemes. In addition, since simulated reflectivity calculations contain considerable uncertainty, a sensitivity study is performed by varying the simulated reflectivity calculation.

2. Data

a. Model forecasts

Two 3-km configurations of the Advanced Research version of the Weather Research and Forecasting Model (ARW-WRF, v3.7.0; [Skamarock et al. 2008](#)) with differing microphysics are evaluated (shown in [Table 1](#)). The WRF forecasts were run operationally in support of the 2015 North Dakota Cloud Modification Project to provide forecasters with daily guidance on convective

TABLE 1. The WRF Model configuration used to generate forecasts over the region depicted in Fig. 1. The cumulus parameterization scheme was only used for the 9- and 27-km domains.

Version	3.7.0
Longwave radiation	RRTM
Shortwave radiation	Dudhia
Land surface model	Noah
Planetary boundary layer	Yonsei University
Cumulus parameterization	Kain–Fritsch
Surface layer	MM5 similarity
Microphysics	WSM6 or Thompson

morphology, intensity, and coverage. In summary, the model microphysics varied between the WRF single-moment 6-class (WSM6; Hong and Lim 2006) and Thompson (Thompson et al. 2004, 2008) schemes, with all other physics held constant. The 3-km model domain where the analysis is performed is predominantly located over western North Dakota, embedded within a 9-km domain that is nested within a 27-km parent domain (Fig. 1). Forecasts had a model top of 50 hPa with 45 vertical levels on a stretched grid with an average spacing of ~ 215 m in the boundary layer and ~ 525 m in the free troposphere. Forecasts were initialized daily at 0000 UTC starting on 1 June 2015 and continuing through 30 September 2015 using the 40-km North American Mesoscale Forecast System (NAM) for the initial and lateral boundary conditions. Forecasts were generated hourly out to 48 h from initialization, but only forecast hours 7–30 were analyzed (further discussed in section 3a). The WSM6 and Thompson schemes were chosen because they are very commonly utilized single-moment microphysical schemes for both operational and research purposes, and many biases in these schemes (particularly precipitation) have been extensively investigated (such as by studies mentioned in section 1). The purpose of using two different microphysical schemes is not to facilitate an intercomparison of schemes, but to highlight the variability present in convective structure and depth within WRF simulations. The size of the 3-km domain and complexity of the microphysical schemes used was also limited by computational power and operational constraints set for forecast completion times.

Determining the reflectivity of hydrometeors in model forecasts is not trivial, especially for microphysical schemes that only retain single-moment (i.e., mass) information, as at a minimum, both hydrometeor concentrations and sizes are required. Both the WSM6 and Thompson schemes provide mixing ratios of cloud water, rain, cloud ice, snow, and graupel hydrometeors. Unlike the single-moment WSM6 scheme, the Thompson scheme is a hybrid scheme that also provides the number concentrations of cloud ice and rain. The simulated

FIG. 1. The WRF Model domains (blue) and 126-km radar range from the KMBX and KBIS WSR-88D where the forecast analysis is performed (red).

reflectivity is calculated internally within WRF (i.e., the “do_radar_ref” option) as in Morrison et al. (2009) and is hereafter referred to as the WRF method. The radar reflectivity factor is determined in a microphysically specific manner, meaning the method utilizes necessary constants and distributions directly from each microphysical scheme. The radar reflectivity factor is determined assuming Rayleigh scattering, which is appropriate for a 10-cm-wavelength radar. In an effort to account for some of the uncertainty in simulated reflectivity calculations, a second commonly used method is utilized and discussed in section 4 to determine the sensitivity to the simulated reflectivity calculations.

b. Radar observations

Radar observations were taken by the 10-cm wavelength WSR-88Ds located at Bismarck (KBIS) and Minot (KMBX), North Dakota. Data for each radar are initially transformed from the native polar coordinates to a uniform grid with a grid spacing of 0.01° (~ 1 km) in the horizontal plane and 1 km in the vertical direction, extending up from 2 to 13 km in height above ground level (AGL; herein, all heights are in reference to AGL) using Radx software (Dixon 2010). Data are mapped out to 126 km from each radar (red circles; Fig. 1), which is the farthest distance where the lowest elevation angle radar beam still samples the primary 2-km analysis height (further discussed in section 3a) assuming standard atmospheric refraction. The relatively short maximum range of 126 km also limits potential issues with radar beam broadening at long distances and losses in detectability. Only radar data nearest to the top of each hour (± 10 min) are used so as to match the hourly output of the model simulations. When data are present from

both radars for a specific hour (and within 6 min of each other), the data are spatially composited using a simple distance-weighted mean. No temporal interpolation is performed when compositing radar observations.

The three-dimensional horizontally polarized radar reflectivity data are used for the evaluation of the model-simulated radar reflectivity. Radar reflectivity is the measure of the backscattered power returned to the receiver by meteorological and nonmeteorological targets and is primarily dependent on the concentration and cross-sectional area of the targets. The observed reflectivity field frequently contains many nonmeteorological artifacts such as ground clutter and biological targets (i.e., birds and bugs) that can have magnitudes equaling or surpassing meteorological echo (e.g., [Steiner and Smith 2002](#); [Lakshmanan et al. 2010](#)). Large “blooms” of reflectivity frequently occur during the evening and overnight hours because of the presence of large amounts of bugs and the enhanced ducting of the radar beam. These artifacts need to be removed prior to the analysis to ensure that only active meteorological targets remain; therefore, radar data undergo simplistic quality control filtering to eliminate nonmeteorological echo.

Dual-polarization data, in particular the cross-correlation coefficient ρ_{HV} and differential reflectivity Z_{DR} , are utilized along with reflectivity to filter out nonmeteorological echo. The ρ_{HV} parameter is a measure of the correlation or similarity between the returned power of the horizontally and vertically polarized pulses. Values of ρ_{HV} close to 1.0 typically indicate meteorological scatters with near-uniform particle phases. Reductions in ρ_{HV} are often seen in mixed-phase regions, reaching values lower than 0.8 in regions of large hail and rain–hail mixtures. Similarly, the melting of snow hydrometers such as in stratiform regions of mesoscale convective systems (MCSs) can also reduce values of ρ_{HV} to <0.7 . Nonmeteorological scatters typically have low ρ_{HV} values of <0.7 ; however, under certain conditions, ρ_{HV} values of nonmeteorological scatters can be high and even approach 1.0 ([Tang et al. 2014](#)). The Z_{DR} is the ratio of the horizontally polarized and vertically polarized backscattered power from the radar pulses. Meteorological values of Z_{DR} near 0 dB typically indicate near-spherical drops or tumbling hail. Values of $Z_{DR} > 2.0$ dB indicate hydrometeors with a large horizontal cross-sectional area such as large drops, while negative values typically indicate vertically oriented ice crystals and conical graupel. More information on the polarimetric variables can be found in [Doviak and Zmíć \(1993\)](#), [Ryzhkov et al. \(2005\)](#), and [Kumjian \(2013a,b,c\)](#).

The quality control steps performed follow many of the individual procedures outlined in [Tang et al. \(2014\)](#). [Tang et al. \(2014\)](#) filter radar echo by performing quality control procedures on radar bins along the radial. Since

the radar data used in this study are gridded and only simplistic control procedures are needed, the quality control procedures are adjusted to work with the gridded dataset. Although the purpose of any quality control procedures is to remove as much nonmeteorological echo as possible while retaining all meteorological echo, for this study more emphasis is placed on removing nonmeteorological echo with reflectivity values that rival those of convection to prevent misidentification of convective cores (following [section 3a](#)). Initially, as in [Tang et al. \(2014\)](#), all pixels with ρ_{HV} of <0.95 are removed unless they are identified as being possible hail cores or regions of melting. Hail cores are identified as pixels containing reflectivity ≥ 45 dBZ and 18-dBZ echo tops > 8.0 km. While [Tang et al. \(2014\)](#) correlated reductions in ρ_{HV} with temperature data to identify the melting layer, in this study the reductions in ρ_{HV} are correlated with enhanced reflectivity regions near the height of the typical melting level to avoid including additional uncertainty associated with temperature data from soundings (such as spatial and temporal representativeness). The melting level is identified as pixels containing $\rho_{HV} < 0.7$ and reflectivity > 30 dBZ in the lowest 5 km, where reflectivity data must also be present in at least three layers of the column, stretching from the surface to 5 km. Following the ρ_{HV} filter, a Z_{DR} filtering procedure [not in [Tang et al. \(2014\)](#)] is included to remove any nonmeteorological pixel clusters that remained. Any pixel containing very high values of $Z_{DR} > 5$ dB coupled with low values of reflectivity < 10 dBZ are removed because these high- Z_{DR} signals are typically present from biological scatters such as insects (e.g., [Browning et al. 2011](#)). Meteorological echoes with Z_{DR} of > 5 dB are largely found in regions of heavy rain or regions containing large drops, which will typically have reflectivity values > 10 dBZ. Last, additional reflectivity filters are applied to the data. Columns are filtered if the rate of change of reflectivity in the lowest two heights is > 50 dBZ or the reflectivity of the lowest level pixel is > 30 dBZ with no echo above. These reflectivity filters enable the removal of any ground clutter with strong reflectivity signals. Once all filters are applied, a nearest-neighborhood method is employed to filter out clutter that contains high ρ_{HV} and that was not removed by the ρ_{HV} or reflectivity filters. For a given pixel containing reflectivity data, if more than one-half of the neighboring pixels have no or missing reflectivity data, the pixel is filtered and any remaining echoes composed of fewer than three pixels (i.e., 3 km^2) are removed [similar to the speckle filter in [Tang et al. \(2014\)](#)]. After the data have been filtered, any pixel-sized gaps created by the initial ρ_{HV} filter that are enclosed by meteorological echo are returned to their original reflectivity value. Examples of

FIG. 2. The (left) raw and (right) quality-controlled composite reflectivity fields as observed by the KBIS radar at (a),(b) 2305 UTC 1 Oct and (c),(d) 0417 UTC 22 Sep 2015.

applying the filters are shown for composite reflectivity observations taken by KBIS for typical ground clutter present during active convection (Figs. 2a,b) and during the overnight bloom (Figs. 2c,d).

3. Evaluation of model forecasts

a. Comparison procedure

To enable a direct comparison, the radar and model data are evaluated on the same grid. Since the radar grid contains lower horizontal grid spacing as compared with the model grid (1 vs 3 km), the radar domain may contain objects smaller than the grid spacing of the model. Since the model cannot generate objects smaller than 9 km^2 , it would unfairly penalize forecasts when such objects are present in the observations; therefore, the radar data are interpolated to the model grid. Furthermore, while the radar has 1-km vertical grid spacing, the model grid spacing is variable; therefore, the simulated reflectivity field is linearly interpolated in the vertical direction to match the radar vertical grid.

For an unbiased comparison, the model data are further constrained temporally and spatially. While the WRF Model forecasts had no downtime, there were

periods when either the KBIS, KMBX, or both radars were down. Forecasts were evaluated hourly when data were present from either or both of the radars, starting at forecast valid hour 7 and continuing until hour 30 (i.e., 0700–0700 UTC), giving the forecast a 7-h spinup time. Similarly, since the model domain covers more area than is observable by the radars, the forecasts are spatially constrained to the area with radar coverage at that time. Although the radar data are composited between two radars, the “cone of silence” is still present at low levels and will influence the model evaluations. The cone of silence refers to the lack of coverage directly above the radar that results from the WSR-88D sets only being able to scan up to a maximum elevation angle of 19.5° . The area covered by the cone of silence is filtered out to 30 km from both radar locations at each height for both the model data and radar data. The 30-km range is approximately the radius of the cone of silence at the highest evaluation height of 13 km using the elevation angle of 19.5° .

In the subsequent analysis, forecast objects are evaluated in two ways. First, observed and forecast convective objects are identified at each height, ranging from 2 to 13 km. Convective objects are defined as continuous reflectivity regions surpassing 45 dBZ

FIG. 3. The (a) total number, (b) area coverage, and (c) mean object area of ≥ 45 -dBZ objects with height. The blue lines represent forecasts using the WSM6 microphysical scheme, the red lines represent the Thompson microphysical scheme, and the black lines represent radar observations. In (c), the solid lines represent the mean object area and the dashed lines are the mean ± 1 standard deviation (note that the dashed lines are not plotted when negative).

(e.g., Caine et al. 2013). This first method treats each altitude level separately and there is no check on vertical extent of a particular high-reflectivity area. The second method includes vertical extent and focuses on convective cores. Convective cores are identified as ≥ 45 dBZ at the 2-km height; this area is then extended across all heights to form a column. Note that while the 1-km height is ideal because of the proximity of echo to the surface and frequent use by forecasters, the 2-km height is chosen to expand the area of analysis. The 1-km height limits the horizontal area of analysis enough to produce detrimental effects on the analysis. For example, even small spatial offsets in forecast convection could locate the convection outside the radar range, reducing the sample size.

b. Forecast assessment

To assess the general performance of forecasts in terms of predictability of ≥ 45 -dBZ objects, the number of objects, object areal coverage, and mean object sizes with height are aggregated across the entire forecast period and are presented in Fig. 3. At the base 2-km height, both the WSM6 and Thompson forecasts perform well when compared with observations, generating only 2.9% and 9.1% more objects, respectively (Fig. 3a). Above 2 km, the differences become much more pronounced. At the typical melting-level height (i.e., 3 km), the reflectivity magnitudes are expected to increase because of the bright band. While the bias in overproducing objects increases to 23.4% for Thompson forecasts at the melting level, WSM6 forecasts underpredict object counts by 17.7%. Between 5 and 10 km, both schemes consistently contain almost double the number of observed objects, with Thompson forecasts overpredicting objects at all heights.

The profile of the total areal coverage of objects (Fig. 3b) follows a similar trend as in Fig. 3a; however, the overprediction of areal coverage is more pronounced than the number of objects. In general, other than the decrease in area at the melting level for WSM6 forecasts (discussed further in section 4), both forecasts considerably overpredict the area covered by objects through all heights. At the 2-km height, WSM6- and Thompson-forecast objects cover 91% and 51.7% more area than the observations, respectively. This overprediction of areal coverage is in part due to more forecast objects but also results from individual objects being too large (Fig. 3c). For Thompson forecasts, both the overprediction of object counts and larger mean object sizes contribute to the bias in areal coverage. For WSM6, the bias in areal coverage in the lower troposphere is primarily due to overprediction of object sizes, as the number of objects is predicted well. In the upper troposphere, the WSM6 object sizes are consistent with observations (especially above 8 km), but the object counts are too high. This indicates that the primary cause of the areal coverage bias transitions from object size to object count as one goes to higher altitudes.

To investigate the internal structure of convection, contoured frequency-by-altitude diagrams (CFADs; Yuter and Houze 1995) are generated for the entire forecast period. CFADs of the reflectivity field with height are found by identifying the convective cores at 2-km height and mapping the area covered by the cores vertically; therefore, the reflectivity of each grid square with height within the column of the 2-km core is included. It is apparent from the CFADs that the forecasts contain a wider spread of reflectivity values within convective cores (Fig. 4). The WSM6 CFAD also contains a different slope

FIG. 4. CFADs of reflectivity within convective cores identified at the 2-km height for (a) radar observations, (b) WSM6 forecasts, and (d) Thompson forecasts across the entire analysis period. The differences between CFADs are shown for (c) WSM6 and radar observations and (e) Thompson and radar observations, where the blue shading indicates underrepresentation by the model and red shading indicates overrepresentation by the model. The observed and simulated reflectivity values are binned in 5-dBZ bins.

in maximum frequencies than observations (Figs. 4a,b). The frequencies of ≥ 45 -dBZ reflectivity values within WSM6 convection remain relatively similar between 2 and 8 km and do not decrease with height as in the observations. In directly comparing the frequency distributions, it is seen that WSM6-simulated convection is frequently too weak by 5–10 dBZ up to 31% of the time below 5 km; however, simulated convection is generally too intense between 5 and 8 km 26% of the time (Fig. 4c). WSM6 is missing the most intense (i.e., >55 dBZ) convection. The Thompson CFAD slope and reflectivity distribution are more similar to the observations but contain notably more spread than the observations (Figs. 4a,d). The wider distribution is caused by the Thompson forecasts containing more intense convection than the observations throughout all heights (Fig. 4e). Thompson and WSM6 perform differently near the melting level, where WSM6 forecasts are frequently too weak by at least 10 dBZ 33% of the time and Thompson forecasts are too strong by at least 10 dBZ 26% of the time (Figs. 4c,e).

To gain additional insight, the likelihood of individual convective cores surpassing certain depths and intensities

is analyzed (Fig. 5). Of all convective cores at 2 km, 20%, 35.7%, and 44.6% of cores retain reflectivity magnitudes above 45 dBZ at 6 km and 3.3%, 5%, and 6.9% reach 10 km for the observations, WSM6, and Thompson, respectively. These results show that the forecasts contain more cores that extend deeper into the troposphere than the observations, with Thompson forecasts containing more than double the number of cores reaching 6 and 10 km. At the 2-km height, only 11.8% of observed cores contain reflectivity values surpassing 55 dBZ and 1.6% surpass 60 dBZ (Fig. 5b), while only a few cores surpassed 65 dBZ and none were observed above 70 dBZ. WSM6 has only 2.8% of cores reaching 55 dBZ, and no cores reach 60 dBZ, indicating the forecast convective cores are too weak (Fig. 5a). Conversely, Thompson convective cores are overly intense and have 52.5% and 30.8% of cores reaching 55 and 60 dBZ, respectively, with a few cores even surpassing 75 dBZ (Fig. 5c). The same trends are visible throughout the vertical plane. To highlight further the intensity biases, the profiles are normalized by the number of identified cores (≥ 45 -dBZ objects) at each height and are shown in Figs. 5d–f. In total, while 4% of

FIG. 5. The frequency of convective cores identified at 2-km height surpassing certain reflectivity thresholds with height for (a) WSM6 forecasts, (b) radar observations, and (c) Thompson forecasts. The percentages of objects that surpass certain reflectivity thresholds normalized by the number of convective cores identified at each height are shown for (d) WSM6 forecasts, (e) radar observations, and (f) Thompson forecasts.

all observed cores surpass 55 dBZ at 10 km, none surpass 50 dBZ in the WSM6 forecasts, and 2% of the Thompson-forecast cores manage to surpass 65 dBZ at the same height (Figs. 5d–f). The highest-reflectivity regions in Thompson extend far into the upper troposphere, while in WSM6 the intensity rapidly decreases with height. To summarize, while both microphysical schemes generate more convective cores with height, the strongest cores are too weak in WSM6 and too strong in Thompson.

4. Sensitivity to simulated reflectivity calculation

Since there is a great deal of uncertainty in deriving the simulated reflectivity field, a second, commonly utilized method is used to determine the sensitivity of the results to reflectivity calculation. For the second method, the radar reflectivity factor is computed following the procedure outlined in Koch et al. (2005, hereinafter referred to as K05), where the radar reflectivity factor is determined using fixed assumptions about the distributions of hydrometeor size and shape for rain, snow, and graupel and primarily relies on their mixing ratios. The K05 method was previously used by a multitude of studies (e.g., Kain et al. 2008; Weisman et al. 2008; Kain et al. 2010; Clark et al. 2012; Pinto et al. 2015), and versions of this method are found in many

widely used postprocessing packages such as Read/Interpolate/Plot (RIP), version 4, and NCAR Command Language (NCL). The K05 method constants (e.g., graupel density) were adjusted to match the constants within the WSM6 and Thompson schemes during calculation.

The aggregated evaluation of objects across the forecast period using the K05 method is presented in Fig. 6. While WSM6 had nearly identical object counts as the observations at 2 km with the WRF method (dashed lines), the K05 method (solid lines) contains 45% more objects, which nearly doubles the areal coverage bias that was already present with the WRF method. At the 3-km melting level height, the underforecasting of object counts and areal coverage found with the WRF method is no longer present, with object counts nearly doubling and the areal coverage being over 3 times the observations with the K05 method. The increase in object counts and areal coverage with the K05 method (relative to the WRF method) is visible across all heights for WSM6. Contrary to the WSM6 K05 results, the Thompson biases in overpredicting the number and areal coverage of objects that were present with the WRF method are similar or substantially reduced throughout all heights with the K05 method. The object count bias is reduced by around

FIG. 6. As in Fig. 3, but using the K05 method. The dashed lines in (a) and (b) present the WRF method for reference (from Fig. 3).

one-half above the melting-level height by using the K05 method; however, the greatest differences are in the areal coverage values. For example, for Thompson forecasts at 4 km, the WRF method areal coverage was 160% greater than the observations, which is reduced to only 25% with the K05 method (Fig. 6b). While Thompson mean object sizes were typically around 30 km² larger than the observations with the WRF method, the K05 method generates objects that have very similar mean sizes (and standard deviations) to

the observations, from 2 up to 9 km. For many heights and parameters, the variability in switching simulated reflectivity calculations results in equal if not larger differences than by retaining the same calculation but switching microphysical schemes.

Analysis of the vertical reflectivity distribution within convective cores using the K05 method reveals that WSM6 CFADs have only minor differences between the two methods (cf. Figs. 7b and 4b). The 5–10-dBZ weak bias at the height of the melting level visible in the WRF

FIG. 7. As in Fig. 4, but using the K05-simulated reflectivity calculation.

FIG. 8. As in Fig. 5, but using the K05-simulated reflectivity calculation.

method is not present using the K05 method and only a slight increase in the strong bias is visible above 8 km (Fig. 7c). The Thompson CFAD changed more noticeably and contains less spread with the K05 method (cf. Figs. 7d and 4d), resulting in the forecast CFAD looking more similar to the observed CFAD. The reduced spread in the reflectivity distributions is due to a lack of the positive (high) intensity bias that was present in the WRF method (Fig. 7e). These reductions in core intensity with the K05 method are also visible in Fig. 8. The high bias in the frequency of WSM6 convective cores reaching higher altitudes is relatively unchanged between the K05 and WRF methods, but the weak bias in reflectivity magnitudes is lessened (Fig. 8a). The amount of WSM6 cores surpassing 50 dBZ at the 2-km height increases from 24.5% in the WRF method to 42.2% in the K05 method, better matching the observations; however, the greatest changes in intensity are limited to below the melting level as the K05 and WRF profiles have marginal differences aloft (Figs. 8a,d). The high bias in Thompson core depths improves, reducing from 44.6% of cores reaching 6 km to 28% (as compared with 20% in the observations). The difference is further improved at 10 km, where only 3.5% of the K05 Thompson cores reach 10 km as compared with 6.9% with the WRF method, where the observed frequency is 3.3% (Fig. 8c). The high intensity bias in the Thompson-simulated reflectivity is also improved throughout all

heights, although forecast cores are still too intense, particularly near the surface (Figs. 8c,f). In general, the normalized frequency of Thompson cores matches the observed cores very well above 6 km, except the strongest cores still contain higher reflectivity values than the observations (Fig. 8f).

While the objective evaluation across the entire period shows differences statistically, a subjective analysis is necessary to see how physically consistent storm systems are and to investigate the entirety of the simulated storm structure. A case representative of the biases found in the analysis is displayed for Thompson forecasts in Fig. 9. The 0100 UTC 2 June 2015 (25 h) forecast includes a mesoscale convective system and scattered isolated deep convection. At all heights, it is apparent that high-reflectivity regions cover more area using the WRF method than the K05 method, which agrees with the overall findings that convective cores are larger and more intense when using the WRF method (Figs. 9a,d). The levels of performance between the two methods are most similar at 6 km, especially in coverage (Figs. 9b,e). Although not the main focus of this study, large differences in anvil intensity exist at 10 km, with the K05 anvils being 5 to almost 15 dBZ more intense than the WRF method and seemingly containing larger anvils (when considering >0-dBZ regions; Figs. 9c,f,i).

A cross section through one of the intense storm cells following from point A to B in Fig. 9 reveals major

FIG. 9. The Thompson scheme's simulated reflectivity field on the native model grid using the (top) **K05** and (middle) WRF methods at (a),(d) 2, (b),(e) 6, and (c),(f) 10 km for a 25-h forecast valid at 0100 UTC 2 Jun 2015. (bottom) The absolute difference between the **K05**-simulated and WRF-simulated reflectivity magnitudes (g) in (a) and (d), (h) in (b) and (e), and (i) in (c) and (f), where red and blue denote higher reflectivity magnitudes in the **K05** and WRF methods, respectively.

differences between the two methods (Fig. 10). The differences noted in the statistical analysis between the two reflectivity calculations are clearly visible in the cross sections. The maximum reflectivity in the **K05** convective core is nearly 10 dBZ weaker than in the WRF method core (Figs. 10a,c,e). The general higher intensity of cores using the WRF method results in wider and taller cores as the 45-dBZ isoline (i.e., the definition of a convective core in this study) extends to ~ 11 km (as compared with ~ 9 km in the **K05** method). The same higher intensity associated with the WRF method is seen for shallower convection (e.g., cell at 49.5°N in Figs. 10b,d,f). There is also a complete lack of a stratiform precipitation region in the **K05** method that is present in the WRF method. Based on the reflectivity field, the (>0 dBZ) anvil appears to be longer and wider in the **K05** method (Figs. 10a,c). Interpretation of these simulated reflectivity fields may be misleading, as one may draw the conclusion that there is little snow aggregation present in the anvil of the **K05** storm; however, the opposite conclusion may be

reached when analyzing the storm using the WRF method, yet the snow hydrometeor fields are identical for both (the differences stem from **K05** using the same fixed size distribution regardless of scheme). The reflectivity field generated by the WRF method and associated with the stratiform region covers almost 2 times the area of the corresponding **K05** field (Figs. 10b,d). The differences between the **K05** and WRF methods are likely caused by **K05** assuming a fixed distribution of hydrometeors; however, more detailed analysis investigating individual hydrometeor fields is required to better understand from where the differences stem.

Figure 11 shows the same case as in Figs. 9 and 10, but for the WSM6 scheme using the WRF method to determine the simulated reflectivity field only. While the forecast convection has different characteristics than those depicted in Fig. 9 because of the variation in microphysical schemes, it is easy to identify the corresponding convective elements. The cross section across

FIG. 10. A vertical cross section of the Thompson scheme's simulated reflectivity field on the native model grid using the (top) K05 and (middle) WRF methods following the black lines in Fig. 9 from (a),(c) point A to point B and (b),(d) point C to point D. (bottom) The absolute difference between the K05-simulated and WRF-simulated reflectivity magnitudes (e) in (a) and (c) and (f) in (b) and (d), where red and blue denote higher reflectivity magnitudes in the K05 and WRF methods, respectively.

this storm system reveals an artificial decrease in the reflectivity magnitude at the 3-km melting level by around 10 dBZ (Fig. 11d), which is commonly present in the WRF method for WSM6. This misrepresentation of reflectivity intensity at the melting level causes WSM6 forecasts to contain a sharp decrease in objects and area at that height (Figs. 3a,b) and was depicted well in the aggregated CFADs (Fig. 4c). It was also noted that WSM6 contained more small, cellular elements than Thompson (which was observed throughout the forecast period); however, these cells are typically weaker (i.e., <40 dBZ) and did not have a pronounced impact on the evaluated convective cores. While the cores for this case were visible extending to the same height in the Thompson and WSM6 forecasts, the anvil regions are also much less pronounced in the WSM6 scheme (Figs. 11c,d as compared with Figs. 9 and 10), which is likely caused by the single-moment representation of rain and snow hydrometeors (e.g., Morrison et al. 2009).

5. Summary and conclusions

Four months of daily summertime forecasts generated by two models with differing microphysics (WSM6 and

Thompson) were evaluated to determine the accuracy in depicting the vertical structure and depth of convection. Forecasts were evaluated by comparing the simulated reflectivity field as internally generated within the WRF Model with observations taken by two S-band radars. An analysis of convective objects aggregated over the entire forecast period identified several differences between observed and forecast reflectivity objects. At 2 km both forecasts performed reasonably well; however, very large differences were found at higher altitudes. Forecasts were found to overpredict the number of convective objects above the melting layer. Mean forecast object sizes were also too large, resulting in convective objects that covered double the amount of area relative to observations. Simulated convective cores were also deeper than in the observations, with forecasts generating between 1.4 and 2.2 times as many observed cores reaching 6 and 10 km. While the number, depth, and coverage of simulated convective cores was greater than the observations, the intensities varied according to microphysics scheme. WSM6 cores were generally too weak, and regions of high reflectivity did not extend far enough vertically. Furthermore, the vertical reflectivity distribution did not follow the same slope as

FIG. 11. The WSM6 scheme's (a) 2-, (b) 6-, and (c) 10-km simulated reflectivity fields on the native model grid using the WRF method for a 25-h forecast, valid at 0100 UTC 2 Jun 2015, and (d) the vertical cross section of reflectivity from point A to point B.

the observations. Thompson cores were significantly stronger, with 13.5% of all cores containing reflectivity values that surpassed the highest observed reflectivity value. Thompson CFADs had a similar slope to observations, but they contained considerably more spread. The results showcase the importance of expanding forecast evaluation into multiple levels, as evaluation at one level may lead to conclusions that are not representative for the entire three-dimensional system. Further development and implementation of three-dimensional evaluation techniques are needed to ensure identification, and subsequent improvement, of forecast biases. Future precipitation evaluation studies should consider integrating three-dimensional reflectivity data into their evaluation, because analysis of the convective vertical structure can be useful in helping to identify where biases in the precipitation field originate from.

To assess the uncertainty introduced by variations in reflectivity calculation, a sensitivity study was performed comparing the WRF-simulated reflectivity field against a commonly utilized method (the K05 method). Although the results using both methods led to the same overall conclusions, there were a number of major

differences between the two methods. For some analyses, the differences introduced by changing the reflectivity calculation method were of the same magnitude as the differences generated by changing microphysical schemes. While the simulated reflectivity field provides a way to evaluate model forecasts in a three-dimensional manner and can reveal a plethora of information about convective processes, determining the radar reflectivity factor from model simulations contains substantial uncertainty, especially for single-moment schemes. Both forecasters and researchers should use the simulated reflectivity field as general guidance but be cautious when relying on it for specific storm attributes. To alleviate some of the uncertainties, development and use of forward radar simulators may produce simulated variables that are more equivalent to those observable by radar, thereby allowing a more direct storm comparison.

Acknowledgments. We thank three anonymous reviewers for providing constructive comments and suggestions on the paper. The author also thanks Tara Jensen for her helpful insights and discussions.

Support for this project was provided by the National Science Foundation (NSF) under Grant AGS-1432930 and the Developmental Testbed Center (DTC). The DTC Visitor Program is funded by the National Oceanic and Atmospheric Administration, the National Center for Atmospheric Research, and the NSF. The support to generate the real-time forecasts was provided by the North Dakota Atmospheric Resource Board. Radar data were provided by the National Centers for Environmental Information and processed using the Radx software package.

REFERENCES

- Adler, R. F., and R. A. Mack, 1984: Thunderstorm cloud height-rainfall rate relations for use with satellite rainfall estimation techniques. *J. Climate Appl. Meteor.*, **23**, 280–296, [https://doi.org/10.1175/1520-0450\(1984\)023<0280:TCHRRF>2.0.CO;2](https://doi.org/10.1175/1520-0450(1984)023<0280:TCHRRF>2.0.CO;2).
- Anderson, J. G., D. M. Wilmouth, J. B. Smith, and D. S. Sayres, 2012: UV dosage levels in summer: Increased risk of ozone loss from convectively injected water vapor. *Science*, **337**, 835–839, <https://doi.org/10.1126/science.1222978>.
- Baldwin, M. E., and J. S. Kain, 2006: Sensitivity of several performance measures to displacement error, bias, and event frequency. *Wea. Forecasting*, **21**, 636–648, <https://doi.org/10.1175/WAF933.1>.
- Barber, K. A., G. L. Mullendore, and M. J. Alexander, 2018: Out-of-cloud convective turbulence: Estimation method and impacts of model resolution. *J. Appl. Meteor. Climatol.*, **57**, 121–136, <https://doi.org/10.1175/JAMC-D-17-0174.1>.
- Barth, M. C., S.-W. Kim, W. C. Skamarock, A. L. Stuart, K. E. Pickering, and L. E. Ott, 2007: Simulations of the redistribution of formaldehyde, formic acid, and peroxides in the 10 July 1996 Stratospheric-Tropospheric Experiment: Radiation, Aerosols, and Ozone deep convection storm. *J. Geophys. Res.*, **112**, D13310, <https://doi.org/10.1029/2006JD008046>.
- Browning, K. A., J. C. Nicol, J. H. Marsham, P. Rogberg, and E. G. Norton, 2011: Layers of insect echoes near a thunderstorm and implications for the interpretation of radar data in terms of airflow. *Quart. J. Roy. Meteor. Soc.*, **137**, 723–735, <https://doi.org/10.1002/qj.800>.
- Burghardt, B. J., C. Evans, and P. J. Roebber, 2014: Assessing the predictability of convection initiation in the High Plains using an object-based approach. *Wea. Forecasting*, **29**, 403–418, <https://doi.org/10.1175/WAF-D-13-00089.1>.
- Cai, H., and R. E. Dumais, 2015: Object-based evaluation of a numerical weather prediction model's performance through forecast storm characteristic analysis. *Wea. Forecasting*, **30**, 1451–1468, <https://doi.org/10.1175/WAF-D-15-0008.1>.
- Caine, S., T. P. Lane, P. T. May, C. Jakob, S. T. Siems, M. J. Manton, and J. Pinto, 2013: Statistical assessment of tropical convection-permitting model simulations using a cell-tracking algorithm. *Mon. Wea. Rev.*, **141**, 557–581, <https://doi.org/10.1175/MWR-D-11-00274.1>.
- Clark, A. J., and Coauthors, 2012: An overview of the 2010 Hazardous Weather Testbed Experimental Forecast Program Spring Experiment. *Bull. Amer. Meteor. Soc.*, **93**, 55–74, <https://doi.org/10.1175/BAMS-D-11-00040.1>.
- , R. G. Bullock, T. L. Jensen, M. Xue, and F. Kong, 2014: Application of object-based time-domain diagnostics for tracking precipitation systems in convection-allowing models. *Wea. Forecasting*, **29**, 517–542, <https://doi.org/10.1175/WAF-D-13-00098.1>.
- Davis, C., B. Brown, and R. Bullock, 2006: Object-based verification of precipitation forecasts. Part I: Methodology and application to mesoscale rain areas. *Mon. Wea. Rev.*, **134**, 1772–1784, <https://doi.org/10.1175/MWR3145.1>.
- , —, —, and J. Halley-Gotway, 2009: The Method for Object-Based Diagnostic Evaluation (MODE) applied to numerical forecasts from the 2005 NSSL/SPC Spring Program. *Wea. Forecasting*, **24**, 1252–1267, <https://doi.org/10.1175/2009WAF2222241.1>.
- Dickerson, R. R., and Coauthors, 1987: Thunderstorms: An important mechanism in the transport of air pollutants. *Science*, **235**, 460–465, <https://doi.org/10.1126/science.235.4787.460>.
- Dixon, M., 2010: Radx C++ software package radial radar data. NCAR, https://ral.ucar.edu/projects/titan/docs/radial_formats/radx.
- Donavon, R. A., and K. A. Jungbluth, 2007: Evaluation of a technique for radar identification of large hail across the upper Midwest and central plains of the United States. *Wea. Forecasting*, **22**, 244–254, <https://doi.org/10.1175/WAF1008.1>.
- Doviak, R. J., and D. S. Zrnić, 1993: *Doppler Radar and Weather Observations*. 2nd ed. Dover, 562 pp.
- Ebert, E. E., and J. L. McBride, 2000: Verification of precipitation in weather systems: Determination of systematic errors. *J. Hydrol.*, **239**, 179–202, [https://doi.org/10.1016/S0022-1694\(00\)00343-7](https://doi.org/10.1016/S0022-1694(00)00343-7).
- , and W. A. Gallus Jr., 2009: Toward better understanding of the contiguous rain area (CRA) method for spatial forecast verification. *Wea. Forecasting*, **24**, 1401–1415, <https://doi.org/10.1175/2009WAF2222252.1>.
- Gallus, W. A., Jr., 2010: Application of object-based verification techniques to ensemble precipitation forecasts. *Wea. Forecasting*, **25**, 144–158, <https://doi.org/10.1175/2009WAF2222274.1>.
- Griffin, S. M., J. A. Otkin, C. M. Rozoff, J. M. Sieglaff, L. M. Counce, and C. R. Alexander, 2017: Methods for comparing simulated and observed satellite infrared brightness temperatures and what do they tell us? *Wea. Forecasting*, **32**, 5–25, <https://doi.org/10.1175/WAF-D-16-0098.1>.
- Held, G., 1978: The probability of hail in relation to radar echo heights. *J. Appl. Meteor.*, **17**, 755–762, [https://doi.org/10.1175/1520-0450\(1978\)017<0755:TPOHIR>2.0.CO;2](https://doi.org/10.1175/1520-0450(1978)017<0755:TPOHIR>2.0.CO;2).
- Hong, S.-Y., and J.-O. J. Lim, 2006: The WRF single-moment 6-class microphysics scheme (WSM6). *J. Korean Meteor. Soc.*, **42**, 129–151.
- Houze, R. A., Jr., 1989: Observed structure of mesoscale convective systems and implications for large-scale heating. *Quart. J. Roy. Meteor. Soc.*, **115**, 425–461, <https://doi.org/10.1002/qj.49711548702>.
- Jensen, T. L., and Coauthors, 2010: An overview of the objective evaluation performed during the Hazardous Weather Testbed (HWT) 2010 Spring Experiment. *25th Conf. on Severe Local Storms*, Denver, CO, Amer. Meteor. Soc., 13B.1, https://ams.confex.com/ams/25SLS/techprogram/paper_175848.htm.
- Johnson, A., and X. Wang, 2013: Object-based evaluation of a storm-scale ensemble during the 2009 NOAA Hazardous Weather Testbed Spring Experiment. *Mon. Wea. Rev.*, **141**, 1079–1098, <https://doi.org/10.1175/MWR-D-12-00140.1>.
- , —, F. Kong, and M. Xue, 2013: Object-based evaluation of the impact of horizontal grid spacing on convection-allowing forecasts. *Mon. Wea. Rev.*, **141**, 3413–3425, <https://doi.org/10.1175/MWR-D-13-00027.1>.
- Kain, J. S., and Coauthors, 2008: Some practical considerations regarding horizontal resolution in the first generation of

- operational convection-allowing NWP. *Wea. Forecasting*, **23**, 931–952, <https://doi.org/10.1175/WAF2007106.1>.
- , and Coauthors, 2010: Assessing advances in the assimilation of radar data and other mesoscale observations within a collaborative forecasting–research environment. *Wea. Forecasting*, **25**, 1510–1521, <https://doi.org/10.1175/2010WAF2222405.1>.
- Koch, S. E., B. S. Ferrier, M. T. Stoelinga, E. J. Szoke, S. J. Weiss, and J. S. Kain, 2005: The use of simulated radar reflectivity fields in diagnosis of mesoscale phenomena from high-resolution WRF Model forecasts. *11th Conf. on Mesoscale Processes and 32nd Conf. on Radar Meteorology*, Albuquerque, NM, Amer. Meteor. Soc., J4J.7, <http://ams.confex.com/ams/pdfpapers/97032.pdf>.
- Kumjian, M. R., 2013a: Principles and applications of dual-polarization weather radar. Part I: Description of the polarimetric radar variables. *J. Oper. Meteor.*, **1**, 226–242, <https://doi.org/10.1519/nwajom.2013.0119>.
- , 2013b: Principles and applications of dual-polarization weather radar. Part II: Warm- and cold-season applications. *J. Oper. Meteor.*, **1**, 243–264, <https://doi.org/10.1519/nwajom.2013.0120>.
- , 2013c: Principles and applications of dual-polarization weather radar. Part III: Artifacts. *J. Oper. Meteor.*, **1**, 265–274, <https://doi.org/10.1519/nwajom.2013.0121>.
- Lakshmanan, V., J. Zhang, and K. Howard, 2010: A technique to censor biological echoes in radar reflectivity data. *J. Appl. Meteor. Climatol.*, **49**, 453–462, <https://doi.org/10.1175/2009JAMC2255.1>.
- Lane, T. P., R. D. Sharman, T. L. Clark, and H.-M. Hsu, 2003: An investigation of turbulence generation mechanisms above deep convection. *J. Atmos. Sci.*, **60**, 1297–1321, [https://doi.org/10.1175/1520-0469\(2003\)60<1297:AIOTGM>2.0.CO;2](https://doi.org/10.1175/1520-0469(2003)60<1297:AIOTGM>2.0.CO;2).
- Mass, C. F., D. Ovens, K. Westrick, and B. A. Colle, 2002: Does increasing horizontal resolution produce more skillful forecasts? *Bull. Amer. Meteor. Soc.*, **83**, 407–430, [https://doi.org/10.1175/1520-0477\(2002\)083<0407:DIHRPM>2.3.CO;2](https://doi.org/10.1175/1520-0477(2002)083<0407:DIHRPM>2.3.CO;2).
- Min, K.-H., S. Choo, D. Lee, and G. Lee, 2015: Evaluation of WRF cloud microphysics schemes using radar observations. *Wea. Forecasting*, **30**, 1571–1589, <https://doi.org/10.1175/WAF-D-14-00095.1>.
- Mittermaier, M. P., 2014: A strategy for verifying near-convection-resolving model forecasts at observing sites. *Wea. Forecasting*, **29**, 185–204, <https://doi.org/10.1175/WAF-D-12-00075.1>.
- Morrison, H., G. Thompson, and V. Tatarskii, 2009: Impact of cloud microphysics on the development of trailing stratiform precipitation in a simulated squall line: Comparison of one- and two-moment schemes. *Mon. Wea. Rev.*, **137**, 991–1007, <https://doi.org/10.1175/2008MWR2556.1>.
- Mullendore, G. L., D. R. Durran, and J. R. Holton, 2005: Cross-tropopause tracer transport in midlatitude convection. *J. Geophys. Res.*, **110**, D06113, <https://doi.org/10.1029/2004JD005059>.
- Pessi, A. T., and S. Businger, 2009: Relationships among lightning, precipitation, and hydrometeor characteristics over the North Pacific Ocean. *J. Appl. Meteor. Climatol.*, **48**, 833–848, <https://doi.org/10.1175/2008JAMC1817.1>.
- Pinto, J. O., J. A. Grim, and M. Steiner, 2015: Assessment of the High-Resolution Rapid Refresh model's ability to predict mesoscale convective systems using object-based evaluation. *Wea. Forecasting*, **30**, 892–913, <https://doi.org/10.1175/WAF-D-14-00118.1>.
- Powell, S. W., and R. A. Houze Jr., 2015: Evolution of precipitation and convective echo top heights observed by TRMM radar over the Indian Ocean during DYNAMO. *J. Geophys. Res. Atmos.*, **120**, 3906–3919, <https://doi.org/10.1002/2014JD022934>.
- Rogers, R. F., M. L. Black, S. S. Chen, and R. A. Black, 2007: An evaluation of microphysics fields from mesoscale model simulations of tropical cyclones. Part I: Comparisons with observations. *J. Atmos. Sci.*, **64**, 1811–1834, <https://doi.org/10.1175/JAS3932.1>.
- Ryzhkov, A. V., T. J. Schuur, D. W. Burgess, P. L. Heinselman, S. E. Giangrande, and D. S. Zrnic, 2005: The Joint Polarization Experiment: Polarimetric rainfall measurements and hydro-meteor classification. *Bull. Amer. Meteor. Soc.*, **86**, 809–824, <https://doi.org/10.1175/BAMS-86-6-809>.
- Schwartz, C. S., and Coauthors, 2009: Next-day convection-allowing WRF Model guidance: A second look at 2-km versus 4-km grid spacing. *Mon. Wea. Rev.*, **137**, 3351–3372, <https://doi.org/10.1175/2009MWR2924.1>.
- Skamarock, W., and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN-475+STR, 113 pp., <https://doi.org/10.5065/D68S4MVH>.
- Steiner, M., and J. A. Smith, 2002: Use of three-dimensional reflectivity structure for automated detection and removal of non-precipitating echoes in radar data. *J. Atmos. Oceanic Technol.*, **19**, 673–686, [https://doi.org/10.1175/1520-0426\(2002\)019<0673:UOTDRS>2.0.CO;2](https://doi.org/10.1175/1520-0426(2002)019<0673:UOTDRS>2.0.CO;2).
- Tang, L., J. Zhang, C. Langston, J. Krause, K. Howard, and V. Lakshmanan, 2014: A physically based precipitation–nonprecipitation radar echo classifier using polarimetric and environmental data in a real-time national system. *Wea. Forecasting*, **29**, 1106–1119, <https://doi.org/10.1175/WAF-D-13-00072.1>.
- 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, [https://doi.org/10.1175/1520-0493\(2004\)132<0519:EFOWPU>2.0.CO;2](https://doi.org/10.1175/1520-0493(2004)132<0519:EFOWPU>2.0.CO;2).
- , P. R. Field, R. M. Rasmussen, and W. D. Hall, 2008: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. *Mon. Wea. Rev.*, **136**, 5095–5115, <https://doi.org/10.1175/2008MWR2387.1>.
- Ushio, T., S. J. Heckman, D. J. Boccippio, H. J. Christian, and Z.-I. Kawasaki, 2001: A survey of thunderstorm flash rates compared to cloud top height using TRMM satellite data. *J. Geophys. Res.*, **106**, 24 089–24 095, <https://doi.org/10.1029/2001JD900233>.
- Van Weverberg, K., N. P. M. van Lipzig, and L. Delobbe, 2011: Evaluation of moist processes during intense precipitation in km-scale NWP models using remote sensing and in-situ data: Impact of microphysics size distribution assumptions. *Atmos. Res.*, **99**, 15–38, <https://doi.org/10.1016/j.atmosres.2010.08.017>.
- Weisman, M. L., C. Davis, W. Wang, K. W. Manning, and J. B. Klemp, 2008: Experiences with 0–36-h explicit convective forecasts with the WRF-ARW Model. *Wea. Forecasting*, **23**, 407–437, <https://doi.org/10.1175/2007WAF2007005.1>.
- Yang, Y. H., and P. King, 2010: Investigating the potential of using radar echo reflectivity to nowcast cloud-to-ground lightning initiation over southern Ontario. *Wea. Forecasting*, **25**, 1235–1248, <https://doi.org/10.1175/2010WAF2222387.1>.
- Yuter, S. E., and R. A. Houze Jr., 1995: Three-dimensional kinematic and microphysical evolution of Florida cumulonimbus. Part II: Frequency distributions of vertical velocity, reflectivity, and differential reflectivity. *Mon. Wea. Rev.*, **123**, 1941–1963, [https://doi.org/10.1175/1520-0493\(1995\)123<1941:TDKAME>2.0.CO;2](https://doi.org/10.1175/1520-0493(1995)123<1941:TDKAME>2.0.CO;2).