Validation of simulated hurricane drop size distributions using polarimetric radar

Bonnie R. Brown1, Michael M. Bell1, and Andrew J. Frambach1

1Department of Atmospheric Sciences, University of Hawai‘i at Mānoa, Honolulu, Hawaii, USA

Abstract Recent upgrades to the U.S. radar network now allow for polarimetric measurements of landfalling hurricanes, providing a new data set to validate cloud microphysical parameterizations used in tropical cyclone simulations. Polarimetric radar reflectivity and differential reflectivity simulated by the Weather Research and Forecasting model were compared with real radar observations from 2014 in Hurricanes Arthur and Ana. Six different microphysics parameterizations were tested that were able to capture the major features of both hurricanes, including accurate tracks, precipitation asymmetry, and the approximate intensity of the storms. A high correlation between simulated intensity and rainfall across schemes suggests an intimate link between the latent heating produced by the microphysics and the storm dynamics. Most of the parameterizations produced a higher frequency of larger raindrops than observed. The Thompson aerosol-aware bulk and explicit spectral bin microphysical schemes showed the best fidelity to the observations at a higher computational cost.

1. Introduction

Mesoscale numerical weather prediction models have become an invaluable tool for forecasting tropical cyclones (TCs), in part due to their ability to represent sophisticated microphysical cloud processes through computationally feasible parameterizations. The nonlinear interaction between the microphysical processes and the model dynamics through the release of latent heat makes the parameterization assumptions particularly critical for forecast accuracy [Igel et al., 2015]. Increased sophistication in a microphysics parameterization is accompanied by increased computational expense, but the inclusion of more hydrometeor types, more physical processes, and more degrees of freedom in the hydrometeor size distributions does not automatically lead to forecast improvements. The desired accuracy must be measured not just by traditional forecast metrics such as TC track and intensity but also in the fidelity of the simulated structure with observations. Though recent work has indicated that newer double-moment microphysics schemes can improve TC forecasts [Jin et al., 2014], the work of Fovell et al. [2009], Pattnaik et al. [2011], Penny [2013], and Khain et al. [2015] among others, show that hurricane forecasts are still highly sensitive to the choice of microphysics.

Two of the primary ways to display and validate numerically simulated TC microphysics are through the use of in situ observations and radar reflectivity. In situ observations can provide a direct measurement of the particle sizes and hydrometeor types and have provided valuable information on drop size distribution (DSD), rainfall rates [Jorgensen and Willis, 1982; Tokay et al., 2008; Black and Hallett, 2012], and snow and graupel size distributions [McFarquhar and Black, 2004]; however, in situ measurements are limited in TCs due to sampling limitations from aircraft and surface instrumentation. In contrast, radar reflectivity can reveal the full three-dimensional spatial distribution of hydrometeors but does not give a direct measurement of the DSD. The radar reflectivity is a combined function of both hydrometeor size and concentration and is heavily weighted toward the largest hydrometeors due to the Rayleigh scattering dependence on the sixth power of the particle diameter at centimeter wavelengths. Many studies have used synthetic radar products to validate the eyewall, rainbands and precipitation asymmetries of simulated TCs, with a few performing more detailed verification of the reflectivity [e.g., McFarquhar et al., 2006, 2012; Rogers et al., 2007]. These studies have shown that simulated reflectivity in model simulations is generally higher than observed, but the magnitude of the overestimation depends on the specific microphysics scheme and the nonlinear interaction with the model dynamics.

Additional information can be obtained by using polarimetric radar, which can measure differences in horizontally and vertically polarized backscatter, propagation phase shift, correlation, and depolarization...
The differential reflectivity factor \( Z_{DR} \), defined as the logarithm of the ratio of the returned power from the horizontally and vertically polarized backscatter, is particularly useful for determining the size of liquid raindrops. Since liquid raindrops take the shape of oblate spheroids that increase their oblateness as they grow, an increase in \( Z_{DR} \) implies an increase in the median drop size of the radar volume. Unlike the conventional radar reflectivity factor at horizontal polarization \( Z_{H} \) (hereafter reflectivity), \( Z_{DR} \) is not sensitive to the number concentration and depends primarily on the shape, orientation, and composition of hydrometeors. While the reflectivity and differential reflectivity alone cannot fully retrieve the DSD without further assumptions, their combination can provide a “fingerprint” of the DSD and uniquely characterize the precipitation [Brandes et al., 2004; Kumjian, 2013].

While dual polarization has been used in radar research for several decades [Seliga and Bringi, 1976], there have been very few observations of TCs using this technology until recently. May et al. [2008] presented the first polarimetric observations of a TC from a 5 cm wavelength radar in Australia. Their analysis of Cyclone Ingrid showed an asymmetric precipitation distribution containing rain–hail mixtures and wet graupel in intense eyewall convection. They did not assess the DSD or show observations of \( Z_{DR} \) in that study, however. Polarimetric mobile Doppler radars have collected some datasets in landfalling hurricanes, but the focus of previous studies has been primarily on the wind structure of the boundary layer, not the microphysical characteristics [Kosiba et al., 2013; Kosiba and Wurman, 2014].

The opportunity to obtain new polarimetric measurements in TCs increased dramatically from 2011 to 2013 with the upgrade of the Next Generation Weather Radar network of coastal Weather Surveillance Doppler Radars (WSR-88Ds). Operational measurements of \( Z_{H} \) and \( Z_{DR} \) can now be made for every landfalling hurricane in the U.S., providing a rich new data set that has recently been used to identify birds in the hurricane eye [Van Den Broeke, 2013] and improve TC rainfall estimation [Ryzhkov et al., 2013; Wang et al., 2014].

In the current study we use these new observations to improve our understanding of TC rain DSDs and to validate the parameterizations of microphysical processes used in numerical hurricane simulations. The results of this study show the polarimetric radar characteristics of two hurricanes from the 2014 season, Arthur and Ana, which were observed by coastal radars in the Atlantic and Central Pacific basins, respectively. Comparisons with high-resolution Weather Research and Forecasting (WRF) simulations indicate that the model is able to adequately capture the overall precipitation asymmetry in both cases. A more detailed comparison reveals some fundamental differences in simulated and observed raindrop sizes using single- and double-moment bulk schemes, as well as an explicit spectral bin model. The results are similar for both storms, suggesting that the conclusions may be more generally applicable to other TCs.

A description of the radar data set and numerical simulations is presented in section 2. The polarimetric radar observations and comparisons with simulations are discussed in section 3. A summary of the results and implications for TC forecasts are discussed in the concluding section.

2. Data and Methodology

The current study uses six different microphysical parameterizations available in the Advanced Research Weather Research and Forecasting model (WRF) [Skamarock et al., 2005] version 3.6.1. The quadruply nested simulations, with the finest grid of 2/3 km resolution, were initialized with the Global Forecast System (GFS) analysis using surface and boundary layer physics tailored for the TC environment. Details of the model setup can be found in Text S1 of the supporting information.

The five bulk microphysical parameterizations tested were the WRF 6-class single- and double-moment schemes (WSM6, WDM6) [Hong et al., 2004; Lim and Hong, 2010], the Lin et al. single-moment scheme [Lin et al., 1983], the Morrison double-moment scheme [Morrison et al., 2005, 2009], and the aerosol-aware Thompson double-moment scheme [Thompson and Eidhammer, 2014]. In addition to bulk parameterizations, a spectral bin microphysical scheme was also tested (“fast” SBM) [Lynn et al., 2005]. All of the microphysical schemes tested predict the same set of mass variables (cloud water, ice, rain, snow, and graupel) but vary in the number concentration variables and representation of cloud physical processes. Table 1 shows the prognostic number concentrations used in each parameterization, where cloud condensation nuclei (CCN) and ice nuclei (IN) denote aerosol number concentrations with no corresponding mass variable.

The bulk schemes use gamma functions of the form \( n(D) = N_0 D^\alpha e^{-\lambda D} \) to represent the DSD, where \( n(D) \) denotes the number concentration of drops with diameter \( D \), \( \alpha \) denotes the shape parameter, \( N_0 \) denotes...
Table 1. Summary of WRF Numerical Experiments

<table>
<thead>
<tr>
<th>Microphysics</th>
<th>Prognostic Intensity Number</th>
<th>Intensity Peak</th>
<th>Peak Accumulated Rainfall (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lin</td>
<td>None</td>
<td>4.5/14.2</td>
<td>401/698</td>
</tr>
<tr>
<td>WSM6</td>
<td>None</td>
<td>2.8/10.3</td>
<td>327/606</td>
</tr>
<tr>
<td>WDM6</td>
<td>Cloud, Rain, CCN</td>
<td>1.1/9.4</td>
<td>289/651</td>
</tr>
<tr>
<td>Morrison</td>
<td>Rain, Ice, Snow, Graupel</td>
<td>3.5/9.2</td>
<td>270/574</td>
</tr>
<tr>
<td>Thompson aerosol</td>
<td>Cloud, Ice, Rain, CCN, IN</td>
<td>1.4/8.7</td>
<td>254/485</td>
</tr>
<tr>
<td>“Fast” SBM</td>
<td>Cloud, Rain, Snow, Ice, Graupel</td>
<td>−3.5/1.3</td>
<td>191/368</td>
</tr>
</tbody>
</table>

*Peak accumulated rainfall refers to the maximum value of the accumulated grid-scale precipitation variable (RAINNC) on the outermost domain at the end of the simulation, which represents precipitation produced by the microphysical scheme but not the convective parameterization. Acronyms are defined in the text.*

the intercept parameter, and Λ denotes the slope of the distribution. The bulk schemes used here have a fixed shape parameter of zero yielding an exponential distribution, with the exception of WDM6 which has a shape parameter of one for rain. With more degrees of freedom, the double-moment physics allow both the intercept and slope parameters of the gamma function to vary independently, while the single-moment physics only allows the slope parameter to vary. The SBM predicts the full spectrum of particle sizes through the use of 33 mass doubling size bins, ranging from a radius of 2 μm to 3.25 mm for liquid drops.

To assess the performance of the chosen microphysical parameterizations, Hurricanes Arthur (2014) and Ana (2014) were simulated. The simulations were initialized at 1200 UTC on 3 July 2014 for Arthur and 0000 UTC on 19 October 2014 for Ana, and then run for 12 and 18 h, respectively. Both storms passed near the coast, with Arthur making landfall in the Outer Banks of North Carolina [Berg, 2015], and Ana tracking near the Hawaiian Islands without making landfall [Powell, 2015]. Table 1 shows a summary of the intensity (maximum 10 m s⁻¹ wind speed) biases and root mean square (RMS) errors of each simulation compared to the best track archive, along with the peak value of accumulated rainfall from each microphysical scheme calculated on the outermost domain. The first three hours of each simulation were removed prior to calculating the average intensity bias and RMS errors over the remaining simulation time. There is a strong correlation with the intensity bias and the peak accumulated rainfall, with $R ≈ 0.93$ and 0.84 for Arthur and Ana, respectively. The strong relationship between intensity and rainfall accumulation is likely due to differences in the latent heating produced by each microphysical scheme. There is a tendency for the more complex schemes to produce both lower rainfall and intensity errors, with the Thompson aerosol-aware and SBM schemes having the lowest total RMS errors, and as will be shown, the best match with the radar observations.

Observations of Arthur were taken by dual-polarized, 10 cm wavelength (S band) WSR-88D radars at Wilmington, NC (KLTX) and Morehead City, NC (KMHX). Ana was observed by the radars at Kauai, HI (PHKI) and Molokai, HI (PHMO). Three hours of observations were chosen to capture the storms at their closest passes to either set of radars, ranging from 2100 UTC on 3 July to 0000 UTC on 4 July for Arthur, and from 1430 UTC to 1730 UTC on 19 October for Ana. Quality control was performed on $Z_h$ and $Z_{DR}$ by removing nonmeteorological echoes and calibrating $Z_{DR}$ using physical constraints. Radar volumes were then interpolated and composited on three-dimensional Mercator grids for subsequent analysis. Further details on the radar data processing can be found in Text S2 of the supporting information.

No attempt was made to directly compare the radar data to WRF rainfall accumulation as such a calculation would involve additional assumptions to convert the observed $Z_h$ and $Z_{DR}$ to rainfall rate. Differences in the simulation time and storm accumulation time further preclude direct quantitative assessment of the rainfall. Instead, the WRF output was converted to polarimetric radar variables to facilitate a direct comparison with the observations. Since the full DSD is known in the model but not in the observations, this forward operator approach involves fewer assumptions than an inverse operator for converting the radar observations to a DSD. The WRF model includes a built-in reflectivity operator using the Rayleigh scattering approximation, but it does not include variations in raindrop shape. In this study we calculate simulated polarized radar variables...
Figure 1. (a–c) The radar reflectivity $Z_H$ (dBZ) and (d–f) differential reflectivity $Z_{DR}$ (dB) of Hurricane Arthur at 2230 UTC on 3 July 2014 from the (a, d) KLTX and KMHX radars and (b, e) the 2/3 km domain of a WRF simulation using the aerosol-aware Thompson microphysics scheme and (c, f) SBM microphysics, at an elevation of 2 km. The sharp edge of the real radar data echo at longer range is due to radar range limitations and not a physical distribution of the hydrometeors.

from the model-derived DSD in post processing as in Jung et al. [2010] and Ryzhkov et al. [2011]. Further details on the polarimetric radar simulator can be found in Text S3.

3. Resulting Distribution in $Z_H - Z_{DR}$ Phase Space

All of the WRF simulations recreate the tracks of Arthur and Ana reasonably well and capture the major asymmetries in precipitation (not shown); however, there is variability in the accuracy of the simulated intensity (Table 1, columns 3 and 4). A snapshot of the observed and modeled $Z_H$ and $Z_{DR}$ for Arthur (2014) at 2230 UTC on 3 July is shown in Figure 1. Only the two most accurate (Thompson and SBM) simulations are shown for brevity. Though there are clear differences, the modeled convection has a similar spatial distribution and general magnitudes of both $Z_H$ and $Z_{DR}$ compared to the observed storm, but with a larger eye and eyewall. Banded structures of $Z_H$ and $Z_{DR}$ are simulated in the correct location, for example, in the northeast quadrant of the storm and in a band in the southeast quadrant. The peak $Z_H$ values exceeding 50 dBZ in the Thompson scheme are similar to the observations, while the SBM underestimates the peak $Z_H$ in Arthur’s eyewall and rainbands. The simulations produce $Z_{DR}$ values that are generally comparable to the observations, but both schemes significantly overestimate the peak $Z_{DR}$ compared to the real radar data.

Figure 2 shows that the observed asymmetric structure and size of Ana (2014) at 1615 UTC on 19 October was simulated fairly well in the $Z_H$ field, with the majority of the precipitation on the northeast side of the TC. The Thompson scheme has a very good representation of the asymmetric eyewall, with peak $Z_H$ values similar to those observed. The SBM captures the overall precipitation asymmetry, but the eyewall structure is not as well defined and the peak $Z_H$ is underestimated. The structure of $Z_{DR}$ east and north of the eye was generally well simulated in both schemes (Figures 2c and 2d), though similarly to Arthur the peak values were much higher than observed. The southwest quadrant of Ana was not observed by radar during the simulation period; therefore, it is not possible to evaluate the high values of simulated $Z_{DR}$ southwest of the eye in Figure 2d.

Both the Thompson and SBM simulations exhibit areas with excessively high $Z_{DR}$ that correspond with moderate $Z_H$ (e.g., Figure 2c) that are not found in the radar observations. These regions were found to have a low
Figure 2. As in Figure 1, for Ana at 1630 UTC on 19 October 2014.

droplet number concentration and moderate mixing ratio which result in the slope of the exponential DSD to be too flat in the Thompson scheme. Radar variables calculated using the full DSD from the SBM model showed a similar signature, suggesting that it was not directly a result of the assumed exponential distribution but rather due to physical processes in the model. A similar radar signature was reported in double-moment schemes by Kumjian and Ryzhkov [2012] due to size sorting of hydrometeors; however, deficiencies in the representation of breakup of large drops or overly aggressive growth processes could also contribute to larger than observed drops. The specific microphysical processes responsible for producing this signature in these hurricane simulations require further study.

To further evaluate the simulations quantitatively, a joint probability density function (PDF) was created in $Z_H\text{-}Z_{DR}$ space and normalized by the maximum frequency in each data set. Each point in this phase-space diagram represents a unique DSD through the combination of $Z_H\text{-}Z_{DR}$, thus validating the model's ability to properly represent the raindrop size statistics throughout the storm. Each PDF encompasses the entire innermost domain and is vertically integrated from the surface to approximately 3.3 km elevation to ensure that only liquid water drops are included. The PDF of the observed $Z_H$ and $Z_{DR}$ is characterized by peak frequencies at 30 dBZ and 0.25 dB in Arthur (Figure 3a) and 35 dBZ and 0.75 dB in Ana (Figure 3c). The modal distribution (frequencies greater than 50% [e.g., Hence and Houze, 2011]) extends from 20 dBZ to 40 dBZ in $Z_H$ and from 0 to 1.25 dB in $Z_{DR}$, while less than 2.5% of the peak frequency exceeds 50 dBZ in $Z_H$ and 2 dB in $Z_{DR}$ in each storm.

The single-moment (WSM6 and Lin) and Morrison schemes all exhibit large frequencies of high $Z_H$ and high $Z_{DR}$ that are not seen in the observations. The Lin microphysics scheme was found to have nearly identical PDFs to WSM6 and thus is not shown. The modal distributions for these parameterizations exceed 40 dBZ $Z_H$ and 1.5 dB $Z_{DR}$ for all single-moment simulations (Figures 3b and 3d); the modal distribution exceeds 50 dBZ and 2.0 dB for both simulations using Morrison (Figures 3f and 3h). The narrow range of DSDs in the WSM6 simulations is primarily due to fixed intercept parameter ($N_0$), such that increases in liquid water content automatically lead to increases in the median drop size through the decrease in the slope of the distribution. The PDFs derived from the single-moment and Morrison simulations imply that the rain rate and the drop sizes are both too large compared to the observations [e.g., Kumjian and Prat, 2014].

The WDM6 simulations look similar to the WSM6 but are truncated in $Z_{DR}$ due to the different shape parameter and have higher frequencies of small drops not seen in the observations, consistent with the results of Gao et al. [2011]. The similarity between these two schemes suggests that the rain DSD in WDM6 may still
Figure 3. Joint frequency distributions of $Z_H$ and $Z_{DR}$ for Hurricanes (left two columns) Arthur and (right two columns) Ana (2014) normalized by the maximum frequency in the data set (percent, gray contours at 0.01, 0.1, 1, and 2.5%, black contour at 50%). (a, c) Composite radar observations for Arthur and Ana (below 3.25 km elevation) and simulated observations derived from the WRF model output using (b, d) WSM6, (e, g) WDM6, (f, h) Morrison, (i, k) aerosol-aware Thompson, and (j, l) SBM microphysics options. Statistics from WRF are limited to below approximately 3.3 km. $Z_H$ is binned from $-10$ to 60 dBZ every 2 dBZ and $Z_{DR}$ is binned from $-1$ to 4 dB every 0.2 dB.

be largely determined by the single-moment ice processes aloft. Modal distributions extending to negative values in $Z_H$ seen in WDM6 may not be well observed by the real radar due to lack of sensitivity at longer range. The Lin, WSM6, WDM6, and Morrison schemes all have a high intensity bias compared to the best track observations.

In contrast, the modal distributions of the Thompson and SBM schemes show greater fidelity to the radar observations and are below 40 dBZ $Z_H$ for Arthur (Figures 3i and 3j) and Ana (Figures 3k and 3l). The PDFs for Ana are generally consistent with those from Arthur but show a slightly poorer comparison with a bimodal distribution and a higher frequency of light rain rates and smaller drops than the observed DSDs. The modal distributions of the simulations using the SBM microphysics are contained between 30 and 40 dBZ $Z_H$ and 0 and 1 dB $Z_{DR}$ but encompass a smaller area of the phase space than the observations. Despite capturing the peak frequencies of moderate convection well, the SBM does not reproduce the low frequency of the very strong convection in the upper right corner of the phase space seen in the observations.

The Thompson scheme exhibits a wider distribution of $Z_{DR}$ for a given reflectivity than the other simulations and captures the low-frequency peak values well, further indicating good agreement with the radar observations. However, the Thompson scheme does exhibit high frequencies of unrealistic DSDs with very high $Z_{DR}$ exceeding 2 dB across a broad range of $Z_H$ (5 to 40 dBZ) as was seen in Figure 2e. As noted above, similar DSDs and radar signatures are also found in the SBM but the PDFs indicate that they occur at a much lower frequency than in the Thompson scheme. This region of the phase space occupies less than 0.1% of the normalized frequency in the radar observations, suggesting that it is an unphysical DSD with too many large drops in light rain compared to real hurricanes.

4. Conclusions

In this study we have compared polarimetric radar variables simulated by the WRF numerical weather prediction model with radar observations of Hurricane Arthur in the Atlantic and Ana in the Central Pacific...
from 2014. Six different microphysics parameterizations with varying levels of complexity were tested, and all of the simulations were able to capture the major features of both hurricanes, including accurate tracks, asymmetric distributions of precipitation, and the approximate intensity of the storms. However, all the simulations showed some discrepancies with the polarimetric radar observations, and most had a high bias in both intensity and rainfall compared with observations.

In general, the model simulated drop size distributions contain a higher frequency of large drops than observed when using the Lin, WSM6, WDM6, and Morrison microphysical parameterizations. These single-moment (Lin and WSM6) and double-moment (WDM6 and Morrison) schemes also had the highest rainfall accumulations and highest intensity errors. Simulations using the Thompson aerosol-aware double-moment scheme and the “fast” spectral bin scheme showed the highest fidelity to the radar observations and lowest intensity errors. The Thompson scheme reproduced the peak and modal frequency of the observed polarimetric characteristics reasonably well but did show a significant frequency of differential reflectivity values that were much larger than observed.

The spectral bin simulation resulted in the most accurate intensity and had a lower frequency of high differential reflectivity values in light rain than the Thompson scheme. However, the spectral bin scheme was unable to reproduce the highest values of reflectivity seen in the observations that were associated with strong convection, resulting in the lowest rainfall accumulation of all the schemes. The spectral bin scheme also required nearly 16 times more computational time to run than the baseline WSM6 scheme using an equivalent number of processors and required significant additional storage for the full DSD, therefore making it unfeasible for operational forecasting at this time.

The high correlation between simulated intensity and rainfall in all schemes suggests an intimate link between the latent heating produced by the microphysical scheme and the storm dynamics. Deficiencies in the representation of microphysical processes could account for the differences in the joint probability distributions of reflectivity and differential reflectivity presented here. Since the low-level rain drop size distribution is a result of the microphysical processes aloft (including ice processes) and the nonlinear interaction with the dynamics, it is difficult to directly attribute discrepancies with the observations to a single microphysical process. For example, a high frequency of excessively large raindrops could be the result of insufficient breakup of large drops, excessive vertical motion, overly efficient collision coalescence, or excessive production of large graupel or snow that undergo melting. Size sorting of hydrometeors and the simplified representation of the DSD in the bulk schemes can also result in differences with the observations.

Future work will extend this analysis to ice and mixed-phase regions and will include other polarimetric variables, such as the specific differential phase and correlation coefficient, with the goal of determining the source of the model deviations from the observations. Further study into the microphysical processes and comparison with polarimetric radar data from tropical cyclones will lead to better forecasts and valuable insights into the mechanisms of hurricane intensification and heavy rainfall production.

References


Computational and Information Systems Laboratory (2012), Yellowstone: IBM iDataPlex system (University Community Computing), National Center for Atmospheric Research, Boulder, Colo., Archive Resource Key. [Available at http://n2t.net/ark:/85065/d7wd3xhc.]


Acknowledgments

This study was supported by the Office of Naval Research award N000141410118 and National Science Foundation CAREER award AGS-1349881. Computing resources from NCAR’s Computational and Information Systems Laboratory (CISL) Yellowstone HPC facility were used for selected WRF simulations [Computational and Information Systems Laboratory, 2012], and GFS analyses were obtained from CISL’s Research Data Archive [National Centers for Environmental Prediction et al., 2000]. Level II radar data were provided by the NOAA National Climatic Data Center. Best track data were obtained from the HURDAT2 data set produced by the NOAA’s National Hurricane Center [Landsea et al., 2015]. Thoughtful and constructive comments from Matthew Kumjian and an anonymous reviewer greatly improved this manuscript.


