FORECAST MODEL ACTIVITIES FOR NORTH DAKOTA CLOUD MODIFICATION PROJECT

Gretchen Mullendore and Mariusz Starzec
University of North Dakota
Corresponding Author: Gretchen Mullendore (gretchen@atmos.und.edu)

ABSTRACT: An overview of regional forecast simulations in support of North Dakota Cloud Modification Project (NDCMP) operations is presented. Simulations are run twice daily during the summer season by the University of North Dakota (UND). Object-based verification of observed and simulated radar reflectivity is conducted during the off-season. Verification of seasonal performance allows modifications to forecast simulations focused on improving operations for the North Dakota region. NDCMP simulated forecasts are found to be biased towards too many convective objects and convection that is too intense and too large. Case studies show that a change in microphysical scheme and adjustments to cloud droplet concentration lessen the biases.

1. OVERVIEW

Since 2012, UND has run forecast simulations for the NDCMP. These forecasts are operational during the summer months, 1 June through 30 Sept (or end of flight operations, whichever occurs first). The forecast configurations that run every season vary depending on both the performance of the previous year and the forecaster needs. All forecasts are run using the Weather Research and Forecasting model (WRF; Skamarock et al. 2008). These 48-hour forecasts are initialized using the 40-km NAM forecasts at either 00 or 12 UTC. The coarsest domain is run at 27-km horizontal spacing and then twice nested to a 3-km domain focused on western North Dakota. A 1-km domain was tested, but both qualitative and quantitative verification showed no increase in performance. The vertical spacing is variable, with 45 levels.

In 2012, only one model configuration was run, but subsequent seasons have consisted of between three and five realizations per day. The multiple model configurations form a primitive ensemble modeling system consisting of varied model cores (different WRF versions, from version 3.1.1 to current), varied microphysics schemes (both packaged and in-house modifications), and two different initialization times. The exact configuration of the Linux computer systems running the simulations has varied over the years, but the most recent configuration (running five forecast simulations per day) is as follows: three simulations run on a multicore machine (using 10-12 cores) and two simulations run on a cluster of slower machines (using 30 cores). Careful benchmarking is conducted each season to maximum processor use while also satisfying the clock time limits for a useable forecast. Forecast products are automatically created and uploaded to a webpage (Figure 1; UND 2015) for access by the NDCMP forecasting team. The webpage displays standard forecasting products as well as specific items requested by the forecasting team, such water vapor mixing ratio at 2 km and HAILCAST (Creighton et al. 2014) mean hail diameter.

Various operational challenges have been encountered over the years. Some notable lessons learned are listed below:

- Posting of the 12 UTC forecasts early enough to be useful for same day operations can be challenging. The UND team found that starting the run prior to all the NAM forecast times posting worked best, with a subsequent restart once the NAM forecasts were fully posted.
- The forecast realizations were sometimes crashing on days with particularly strong convection (days of most concern for hail suppression operations). The best solution found was to turn on selective vertical damping in the WRF namelist.

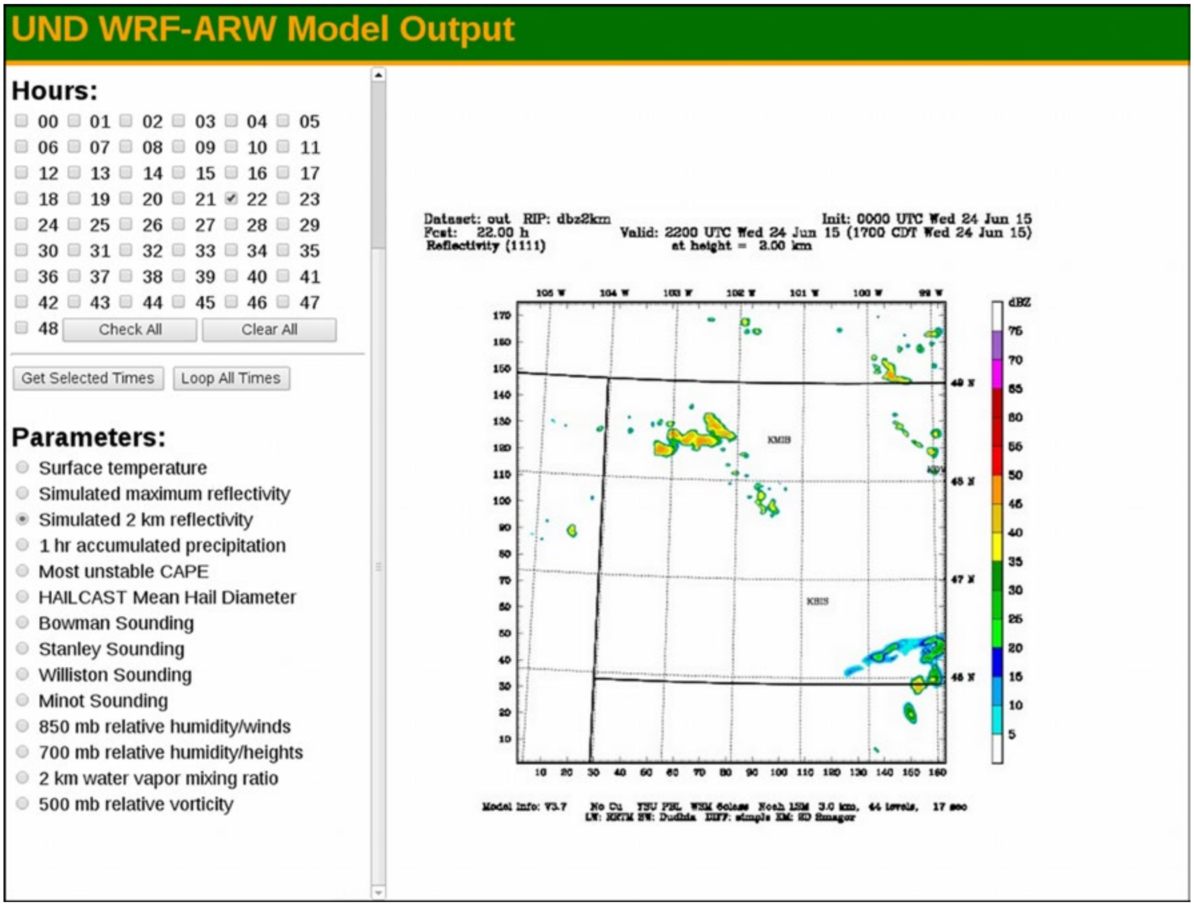

Figure 1: Example screenshot from the forecast simulation webpage.

2. FORECAST VERIFICATION

2.1. Verification Methodology

Forecasts are evaluated by comparing 1 km above ground level (AGL) simulated reflectivity field to the radar reflectivity as measured by C-band Doppler radars. Forecasts are verified following an objective object-based methodology, where storm objects are generated in both forecast and observation reflectivity fields at each valid forecast time (i.e., hourly for this study). Storm objects are discrete regions of reflectivity above a certain magnitude threshold. The number of objects and their corresponding sizes at different reflectivity magnitude thresholds are compared between forecasts and observations to determine if forecasts correctly represent the convective coverage, morphology, and intensity. Forecasts are also assessed on how many hits, false alarms, and misses occur over the forecast period, dependent on whether or not convection was present in both the forecast and observation domains. Forecasts initialized at 00 UTC are evaluated across 21 hours starting at 03 UTC to allow for model spin-up time, and extend until 00 UTC the next day. Since the modeled area is much larger than the area covered by the radar, only model data within the 150 km range of the radars is utilized for the comparison, and any model data outside the radar range is excluded from the analysis. The modeled reflectivity and radar reflectivity are both compared at 3 km horizontal grid spacing.

2.2. Local Verification and Forecast Improvement

Across the NDCMP 2012 and 2013 operational seasons, 3537 times across 186 days are evaluated. Figure 2 shows the number of model objects (solid bars) and radar objects (hashed bars) binned according to object area at different reflectivity thresholds of greater than 5 (green), 30 (yellow), and 45 dBZ (red). Across all simulations, forecasts over-predict the number of convective cells across all size bins for all magnitude thresholds except for the smallest size bin (1-45 km^2). The smallest size bin contains many small radar artifacts that were not filtered out and bias the bin; therefore, it is ignored. The model simulations significantly over-forecast the number of weak (5 dBZ) and strong (45 dBZ) objects. A general improvement in forecasts is visible when increasing the reflectivity threshold from 5 dBZ to 30 dBZ. The improvement indicates that forecasts contain too many weak (5 dBZ) cells. However, all forecasts have less skill at the 45 dBZ threshold, making it evident that the forecasts also over-intensify convection, as a smaller amount of radar objects exist at the higher reflectivity threshold requirements.

A standard hit, miss, and false alarm rate histogram, based on whether any convection is present in either domain, is shown in Figure 3. A hit case means convection exists in both the forecast and observation domains, while a null case signifies that neither domain contains convection. At a reflectivity threshold of 5 dBZ, success rates (hits plus null cases; blue colors in Figure 3) of 60% are visible across all cases. The high miss rate (red) at 5 dBZ and lower false alarm rate is caused by radar artifacts (i.e. clutter, noise) being present in the domain that was visible in the smallest bin in Figure 1. As the reflectivity thresholds increase to 30 and 45 dBZ, the hit rate drops and the null rate increases as expected, since lower reflectivity values (i.e., weaker cells) are now being masked, leading to an increase in success rates. However, the ratio of the number of false alarms and misses to hits increases.

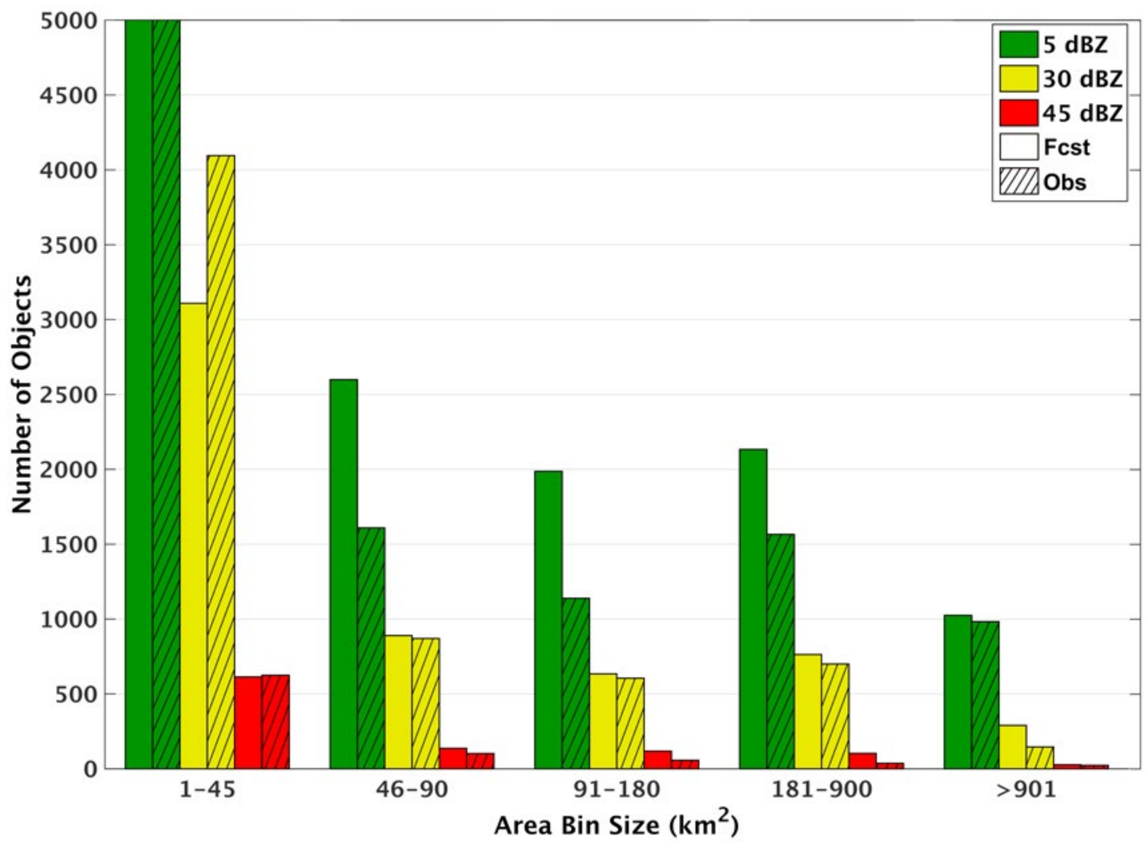

Figure 2: The number of forecasted (solid) and observed (hashed) objects binned according to area across all simulations at reflectivity thresholds of 5 (green), 30 (yellow), and 45 dBZ (red). The vertical axis is limited to 5000 objects in order to preserve plot detail. The actual number of forecasted and observed objects at 5 dBZ is 9313 and 11596, respectively.

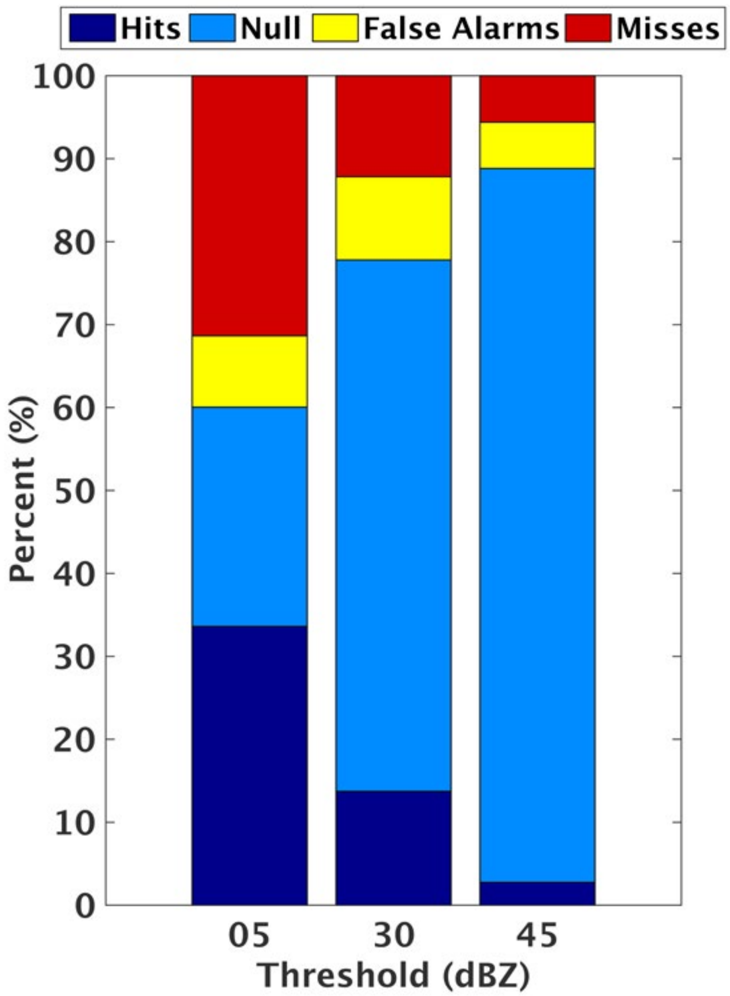

Figure 3: The percentage of cases across all simulations there were either hits (dark blue), null cases (light blue), false alarms (yellow), or misses (red) at differing intensity thresholds of 5, 30, and 45 dBZ.

Forecasts are further evaluated to determine how accurately they predicted the areal coverage of convection at each matching time that was considered a hit. The percentage of the domain covered by objects was calculated for both the radar and model at different intensity thresholds. The absolute difference in areal coverage between radar and model domains is shown plotted in Figure 4. Overall, the forecasts performed relatively well, with 89.9% of forecasts at 5 dBZ being within 10% areal coverage when compared to observations. At the 30 and 45 dBZ thresholds, 96.7% and 100% of all forecasts are within 10% areal coverage of observations, respectively.

In summary, the NDCMP forecasts were found to generate 1) too many convective objects, 2) too much area covered by convection, and 3) over-intensified convection. These biases were subsequently investigated by a sensitivity study performed on various horizontal resolutions, differing microphysics (control microphysics was WSM6; Hong and Lim, 2006), and cloud droplet concentrations (CDC) for several exemplar case days (Starzec, 2014). Results from the sensitivity study suggested that the CDC significantly affected convective forecast skill. The Thompson microphysics scheme (Thompson et al. 2008) with a CDC of 300 per cubic centimeter was found to more accurately represent convective trends than the control scheme (default CDC of 300 cm^{-3}) or the Thompson scheme (default CDC of 100 cm^{-3}). Much of the sensitivities seen for the two microphysics schemes tested here were due to differences in treatment of the autoconversion. Also of note is that Thompson et al. (2008) strongly advised that the CDC be changed from 100 cm^{-3} (representative of clean marine air) to ambient or known concentrations. This updated microphysics was then implemented as one of the members in the next season's ensemble forecast system.

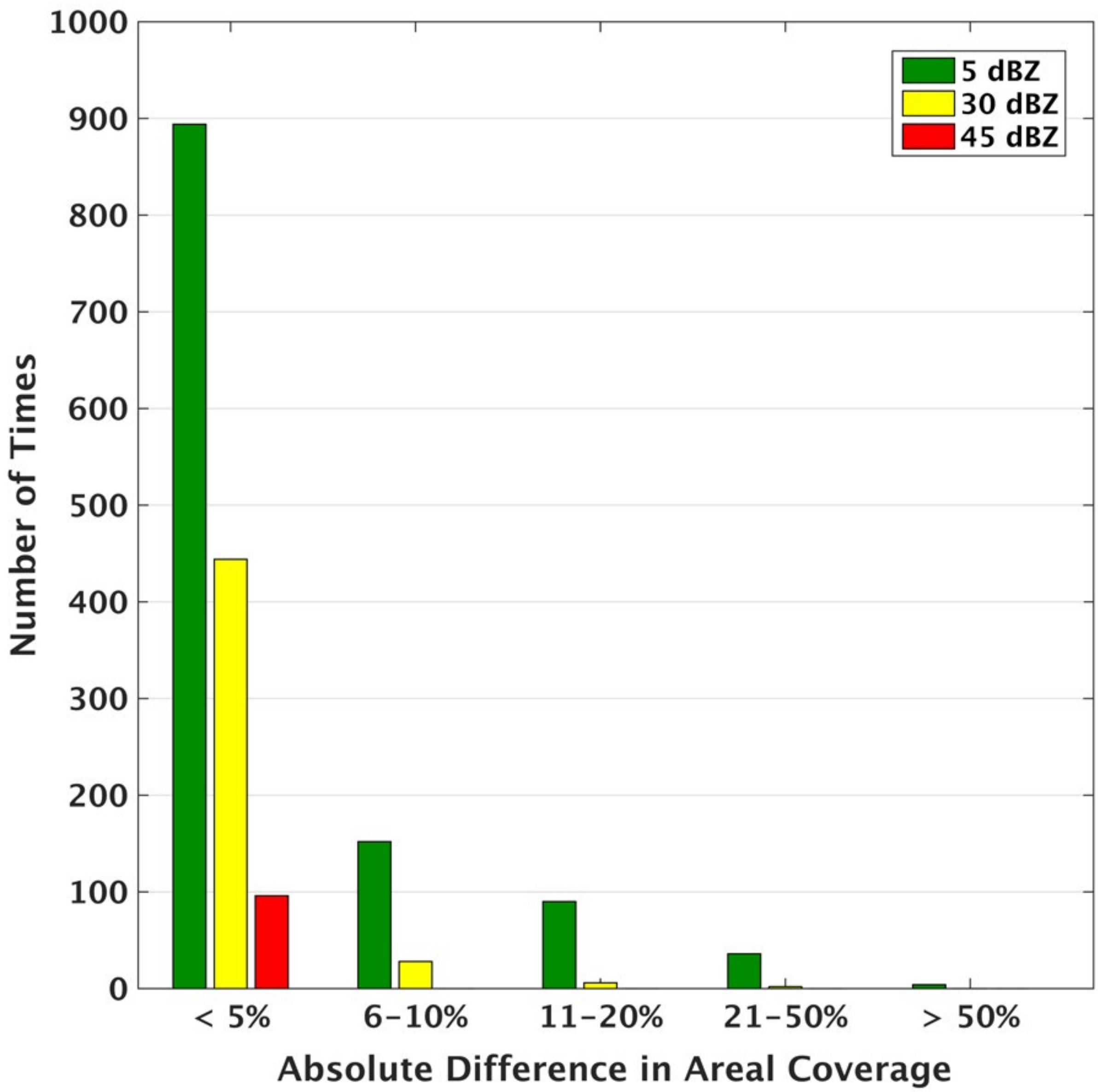

Figure 4: The absolute difference in the percentage of the domain covered by objects in both the observed and model domains at different reflectivity thresholds of 5 (green), 30 (yellow), and 45 dBZ (red).

3. SUMMARY

Regional forecast simulations are run to aid NDCMP forecasters. Originally, these forecasts provided higher resolution forecasts than available operationally. However, now 3-km is not unique. Instead, the primary value is the ability to adapt to the needs of local operations (e.g., HAILCAST) and suggest improvements using focused verification procedures. Verification of seasonal performance allows modifications to simulation setup to provide more accurate forecasts specifically tailored for summertime operations in the North Dakota region.

ACKNOWLEDGEMENTS

The authors wish to acknowledge Darin Langerud and the North Dakota Atmospheric Resource Board (NDARB) for funding this research and the Polarimetric Cloud Analysis and Seeding Test (POLCAST) campaigns, and David Delene for helpful feedback, discussions, and support for this project. The authors would also like to thank the Research Application Laboratory (RAL) visitor program, and many members of both RAL and the Development Testbed Center (DTC) for their valuable suggestions, feedback, and insight. Finally, the authors wish to thank the entire North Dakota Cloud Modification Project team. The UND forecasts constitute only a small part of the NDCMP effort; feedback and support from project forecasters are responsible for making the UND forecasts successful.

REFERENCES

Creighton, G., and Coauthors, 2014: AFWA Diagnostics in WRF. UCAR Note, 17 pp.

Hong, Song-You, and J.-O. J. Lim, 2006: The WRF single-moment micro- physics scheme (WSM6), *J. Korean Meteor. Soc.*, **42**, 129-151.

Skamarock, W. C., and Coauthors, 2008: A Description Of The Advanced Research WRF Version 3. NCAR Technical Note TN-475, NCAR, 113 pp.

Starzec, M., 2014: Forecast Bias Analysis using Object-based Verification of Regional WRF Summertime Convective Forecasts, Thesis in Department of Atmospheric Sciences, University of North Dakota, 132 pp.

Thompson, G., Field, P., R., Rasmussen, R. M., Hall, W., D., 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.

UND 2015: UND WRF-ARW Model Output, http://gretchen.atmos.und.edu/forecast_west/index.xhtml