Accuracy of Locally Forecasted Precipitation as Determined by UND Radar

Mariusz Starzec (mariusz.starzec@my.und.edu)[1], Gretchen Mullendore[1], David Delene[1], Paul Kucera[2]

[1]University of North Dakota, [2]Research Applications Laboratory, National Center for Atmospheric Research

Background

During the Summer of 2010, the Polarimetric Cloud Analysis and Seeding Test 3 (POLCAST 3) field campaign was taking place in order to research the use and effects of hygroscopic seeding flares on convective cells in North Dakota using the University of North Dakota's (UND) Citation Research Aircraft. During the campaign, local Weather and Research Forecasting (WRF) model runs were used to predict the timing, intensity, and distribution of convection and precipitation in the target region. This output information was used ahead of time to plan flights for the Research Aircraft, cloud modification aircraft, and determine whether the team should be put on standby. During the times of predicted precipitation, UND's NorthPol Polarimetric Radar was scanning the region to collect data and to lead the aircraft to convective cells that had a potential for modification.

Objectives

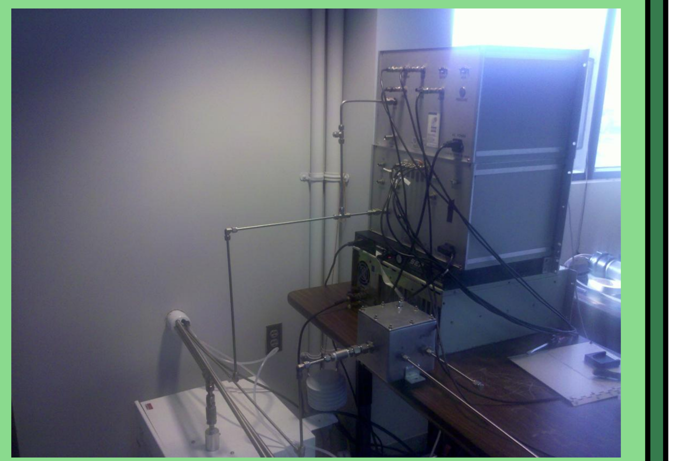
The purpose of this project is to find a method compare forecasted precipitation to observed precipitation, and to effectively compare them. This will be done by comparing the WRF model forecasts to the Radar data. By comparing the forecasted and observed precipitation, the goal is to determine the accuracy of the WRF model. This will also show how reliable the locally run WRF model is when used in deciding whether to prepare the Research Aircraft. Using these results, another goal is to show that traditional point-to-point skill scores should not be used to assess forecasts, since they do not account for changes in location.

UND NorthPol C-band Polarimetric Radar

UND Citation Research Aircraft

Cloud Modification Aircraft

Surface Measurements



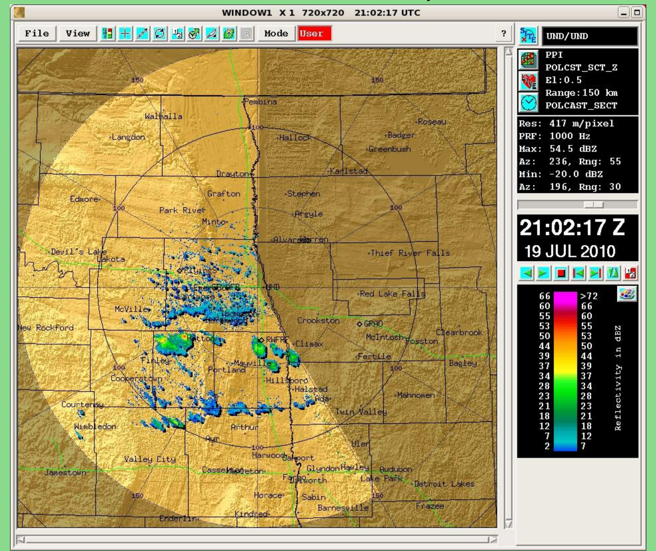
Methodology

Ten test cases were selected to be initially tested, each with a forecast lead time between 14 and 28 hours. Test cases were selected randomly from the subset of cases which contained both observed and forecasted convection. Each case consisted of a single analysis time. Because forecasted results are only archived hourly, the radar data times selected are the closest data to the top of the selected forecast hour. The skill of the forecasted convection was assessed by using traditional point-to-point methods and by using object-based methods. Object-based comparison was done using Developmental Testbed Center's (DTC) Model Evaluation Tools (MET), namely the Method for Object-Based Diagnostic Evaluation (MODE) tool. The MODE tool was chosen as it compares different observed and forecasted cells against each other, as opposed to using traditional verification methods where you compare observations and forecasts point-to-point. It is well known that forecasting the location of convection is difficult and point-to-point verification methods do not account for variations in location.

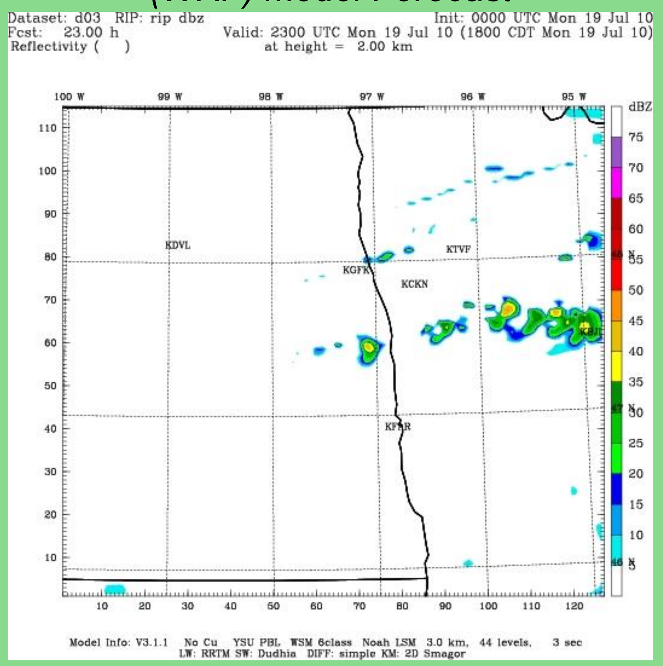
The forecasts and the observations were to be compared using reflectivity values. The first step was to get both the model data and radar data, which were both in different formats, to the format required by the MODE tool. The MODE tool needed all inputs to be in a specific format, which is outputted from the PCP-Combine Tool, another MET tool. However, PCP-Combine only deals with adding or subtracting accumulated precipitation from multiple files into one. Specialized programs were written to complete this conversion step. For the model data, a program was created to calculate reflectivity using the model hydrometeor fields, interpolate it across the same 3-km grid but at one kilometer altitude, and output the data so it exactly matches the output of PCP-Combine. For the radar data, a program was created to take the radar data at one kilometer, interpolate it over the 3-km model grid, and to output it so it matches the output of PCP-Combine. These conversions programs produced the forecast files and observations files in a format that MODE can use, while using reflectivity values and not accumulated precipitation. In addition to the format challenges, while the radar was scanning the region, many times it only scanned sectors and not complete scans, in order to better help guide the aircrafts. Another program was developed to take the radar azimuth data and create polygon files that were used to mask out the regions of that were not scanned and did not contain data. Masking out regions of missing data guaranteed that MODE wouldn't include missing data in its calculations.

After these steps the MODE tool was run using threshold of greater than or equal to 30 dBZ. This threshold was selected since 30 dBZ can be considered as the point where precipitation is almost guaranteed, and is a good indicator of convection. Cells with these reflectivity values were also targeted for modification by the research aircraft during POLCAST3.

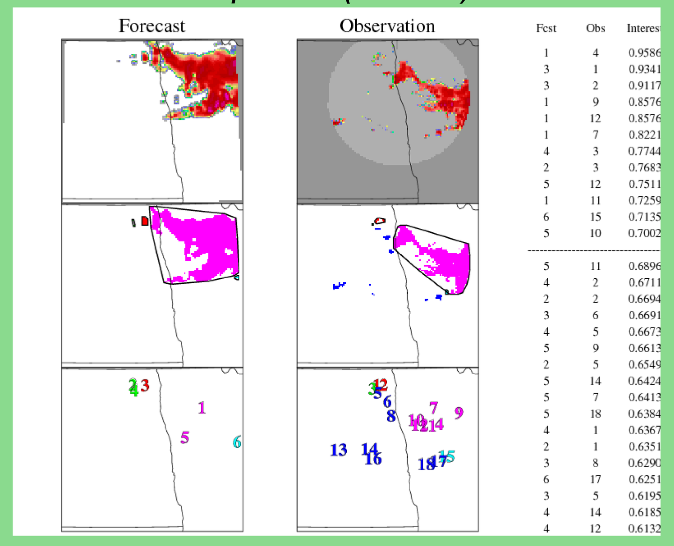
UND's NorthPol Radar Display (Using IRIS Software)



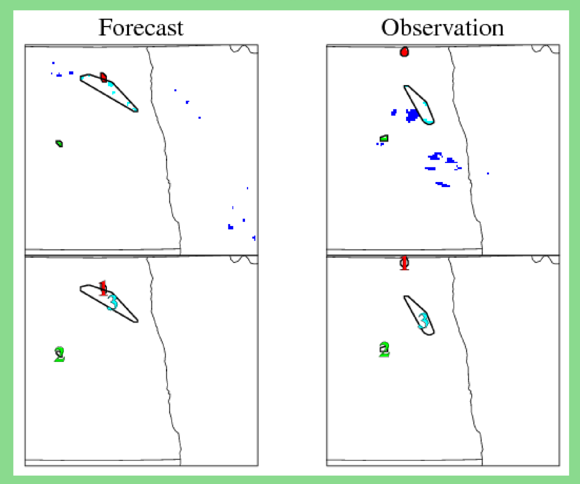
Weather Research and Forecasting (WRF) model Forecast



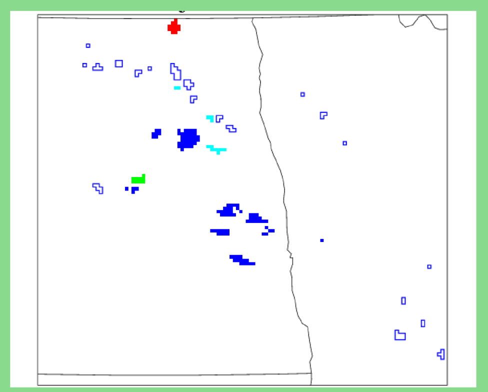
Method for Object-based Diagnostic Development (MODE) Tool



MODE Tool Cluster Object Information



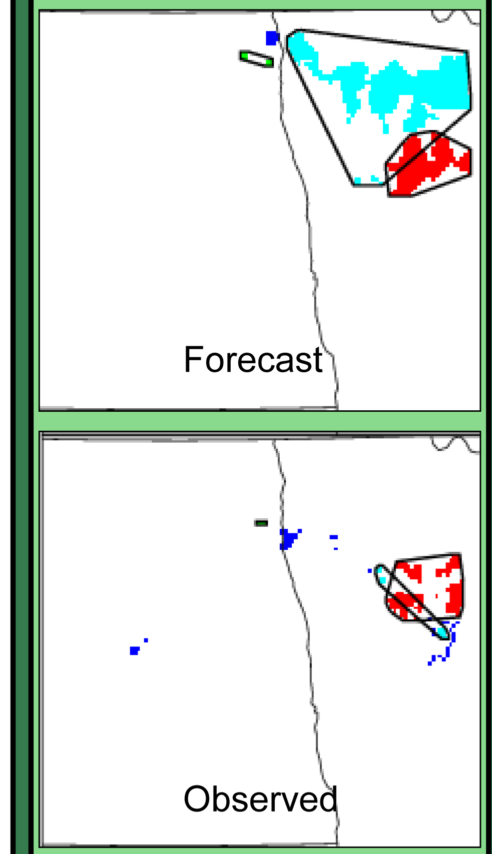
Observation Objects with Forecast Outlines



Analysis

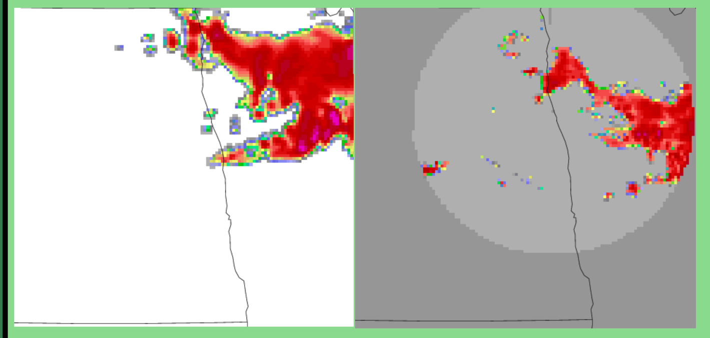
Cell Comparison

Subjective analysis of the MODE tool output showed problems with the way the MODE tool was grouping, clustering, and matching objects (figure to the right). Due to these problems, the clustering and matching features of MODE were not used in the analysis. When MODE defines objects, it calculates the area of each object (cell). Since forecast location is not important at this scale, we decided to compare the count of objects forecasted to the count of objects observed. Each object was put in bins based on area to better visualize the data. Some sample plots of this comparison are displayed as the histograms to the right. By looking at these plots, you can see that the WRF model did fairly well in forecasting the correct count of cells by area. All the test days are included in the large histogram underneath the four sample cases. For most cases and area bins, the difference between the number of forecasted convective objects and the number of observed convective objects was less than three cells. Subjective review of the final test case showed that the poor forecasting results were due to timing of the system.



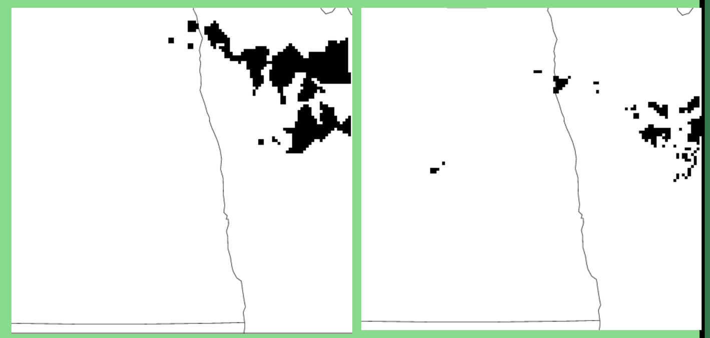
Example of unusual grouping and clustering by MODE.

24 hr Forecast **Observation**

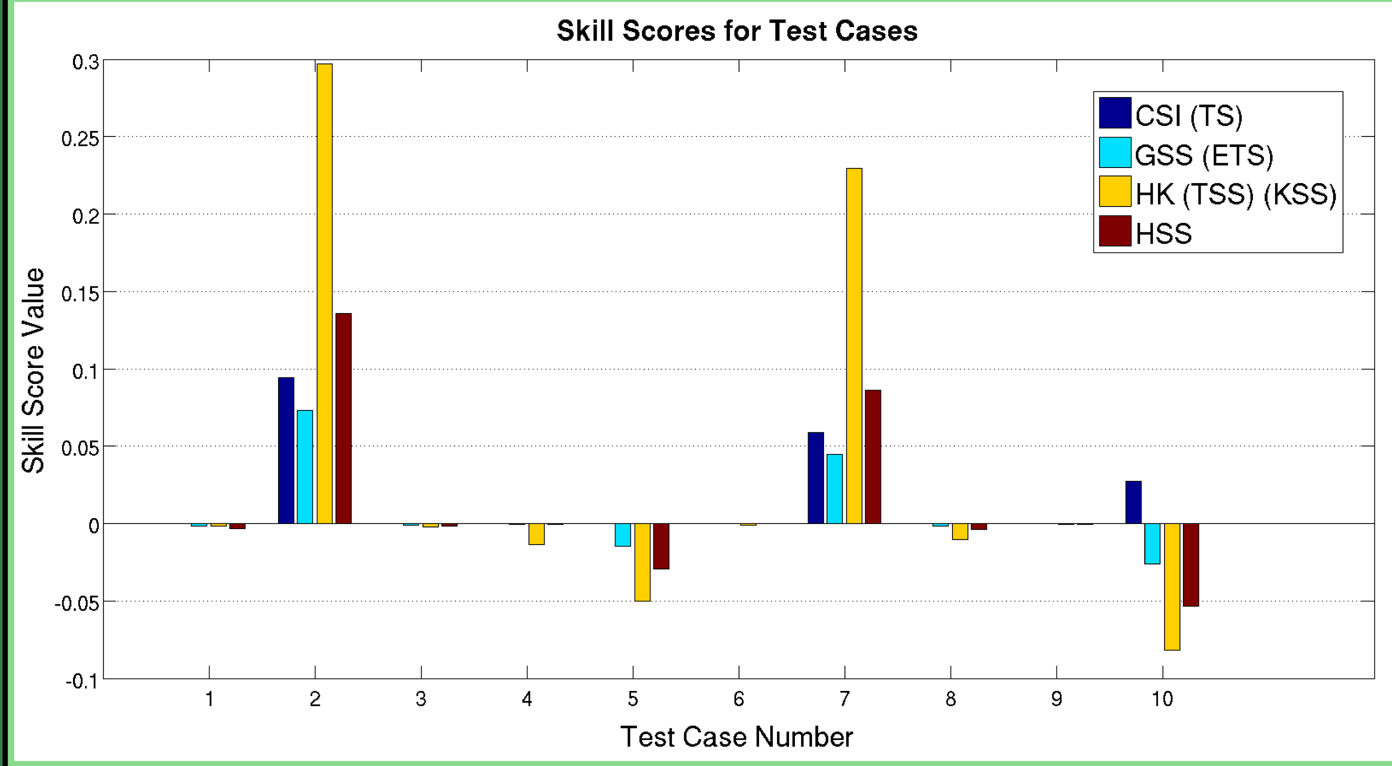


Radar Returns (June 25th, 2010)

24 hr Forecast **Observation**



>30 dBZ Cloud Cells



Histogram plot above shows the calculated traditional point-to-point skill scores for each test case. Skill scores of >0 give more skill to the forecast (score of 1 equals a perfect forecast), scores of 0 give equal skill to the forecast and a random forecast, and scores of <0 state that a random forecast has more skill than what was forecasted.

The chart and formulas on the right show how each skill score is calculated.

Event Forecast	Event Observed		
	Yes	No	Marginal Total
Yes	a	b	a+b
No	c	d	c+d
Marginal Total	a+c	b+d	a+b+c+d = n

$$HSS = 2\,(ad - bc) / [(a+c)(c+d)+(a+b)(b+d)]$$

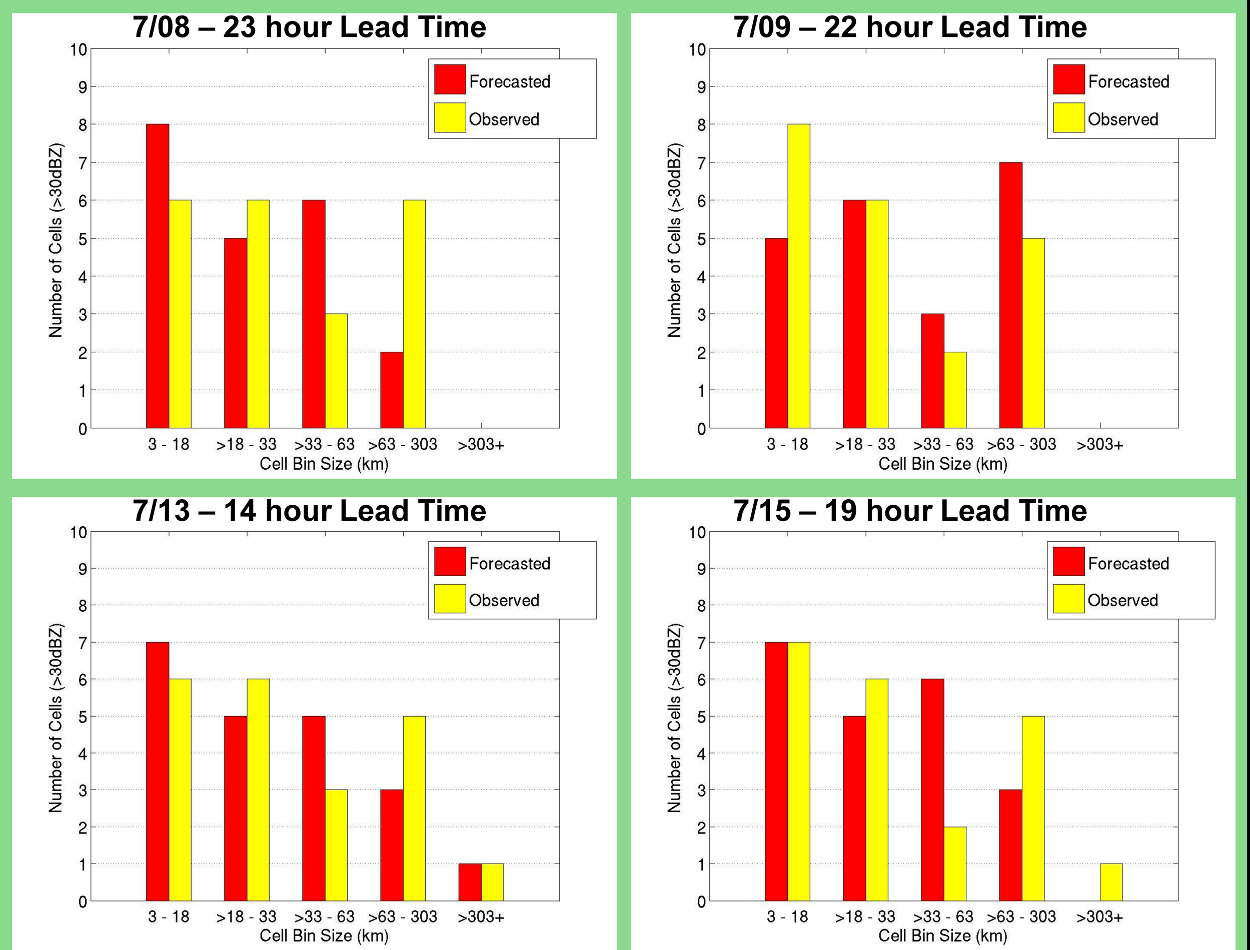
$$HK = KSS = TSS = H - F = (ad - bc) / [(a+c)\,(b+d)]$$

$$GSS = ETS = (a - a_r) / (a+b+c-a_r), \text{ where } a_r = (a+b)\,(a+c) / n$$

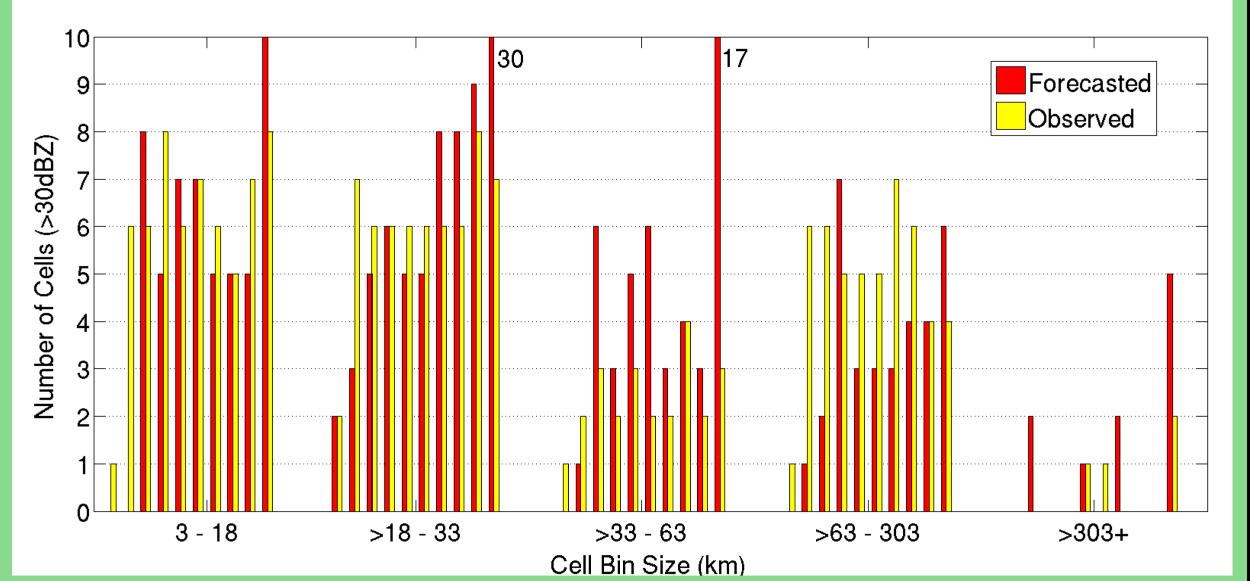
$$CSI = TS = a / (a+b+c)$$

Skill Scores

By looking at the histogram above, we can see that overall the skill scores did not show much skill in the WRF model forecasts. This is simply because if a point of the forecasted grid does not match a point on the observed grid, it instantly counts that point forecast as a miss, even if the location of that point was off by one grid square. Since forecasting the exact location of convection and precipitation is extremely challenging, these skill scores should not be used to assess these types of forecasts. This is why object-based evaluation of precipitation needs to be done to account for the changes in location and, to some degree, shape.



Four sample histograms comparing the forecasts and observations by showing the number of cells in different size ranges that are above the 30 dBZ threshold. By looking at the graphs, you can see that the WRF model did fairly well, generally with about one or two cells difference per bin.



Histogram containing all ten test cases, comparing the forecasted and observed number of cells in different size ranges. The forecast lead times are between 14 and 28 hours, and are from 3 km resolution runs of the WRF model.

Conclusions

Looking at the initial results from the ten test cases we can see that:

- While the predicted precipitation may have not been in the correct location, which is to be expected, the WRF model did fairly well in predicting the distribution and intensity of cells.
- The WRF model 14-hour to 28-hour convective forecasts agree well with observations, and the model can be used to help guide future flights with a high certainty of success.
- Object-based analysis has been shown to be preferable to point-to-point analysis for assessment of convective forecasting.

Future Work

While there is a lot that still needs to be done, and can be done using the data collected, the main goal for the immediate future is to go through all the data collected during POLCAST 3. This includes the 0Z model data and the 12Z model data. Once all the data has been processed, more statistically accurate results will be available to assess the WRF model overall. More data will also allow further investigation as to why the WRF model sometimes excels and other times fails in convective forecasts. The next step would be to create a process that accounts for time lags. There are times when the WRF model correctly simulates the coverage and intensity of a convective system, but the timing is off. Comparisons of the 0Z and 12Z runs can then be compared to show which is more reliable, and whether the 12Z forecast is more accurate as expected.

Acknowledgements: This research has been supported by the North Dakota Atmospheric Resource Board.