Object-based Verification of Regional WRF Simulations of Summertime Convection

Mariusz Starzec[1], Gretchen Mullendore[1], David Delene[1], Paul A. Kucera[2], Darin Langerud[3]

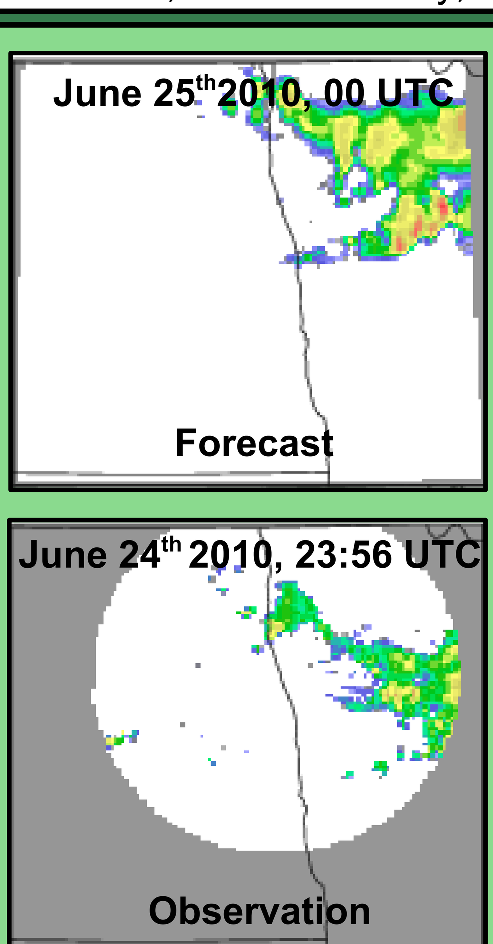
[1]University of North Dakota, [2]Research Applications Laboratory, National Center for Atmospheric Research, [3]North Dakota Atmospheric Resource Board

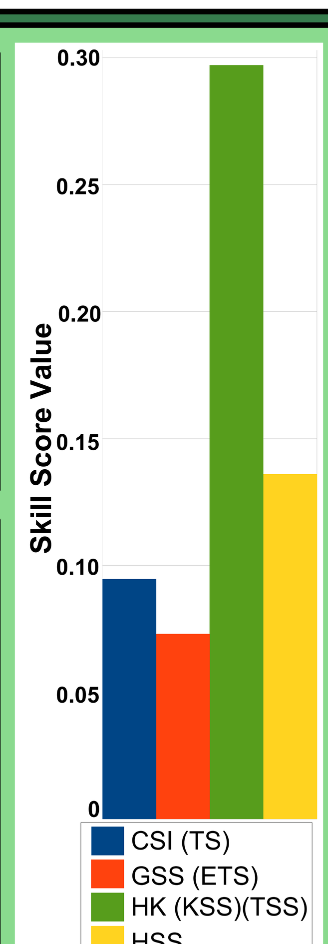
Introduction

As model resolution increases, convection is more explicitly resolved leading to a more detailed and arguably more skillful forecast. To determine whether predictions of convection are actually improving, careful evaluation of forecast skill is needed. Due to the inherent chaotic nature associated with convective initiation, traditional point-to-point verification may not accurately represent the true forecast skill. A slight shift in convective location can result in the forecast being deemed not skillful even if the forecast was subjectively a very accurate representation of the event. Therefore, a method to account for spatial variations is required to make an objective analysis of forecast skill.

Background

The Polarimetric Cloud Analysis and Seeding Test (POLCAST) is an ongoing research project with biennially occurring field campaigns in eastern North Dakota focused on evaluating the effectiveness of hygroscopic seeding flares in summertime convection. Since 2010, local 3-km resolution Weather Research and Forecasting (WRF) model runs have been used to predict the timing, intensity, and distribution of convection in the study region. In 2012, a 1-km WRF run was added to the eastern North Dakota forecast domain. The University of North Dakota's (UND) C-band polarimetric Doppler radar actively scanned the domain during the field campaigns. In addition to the data collected during POLCAST, a 3-km grid spacing model run with a partially nested 1-km grid was used to forecast convection for western North Dakota's ongoing weather modification project. Observations in the western domain were recorded by two C-band dual polarimetric Doppler radars located in Bowman, ND and Stanley, ND.





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

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

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

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

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

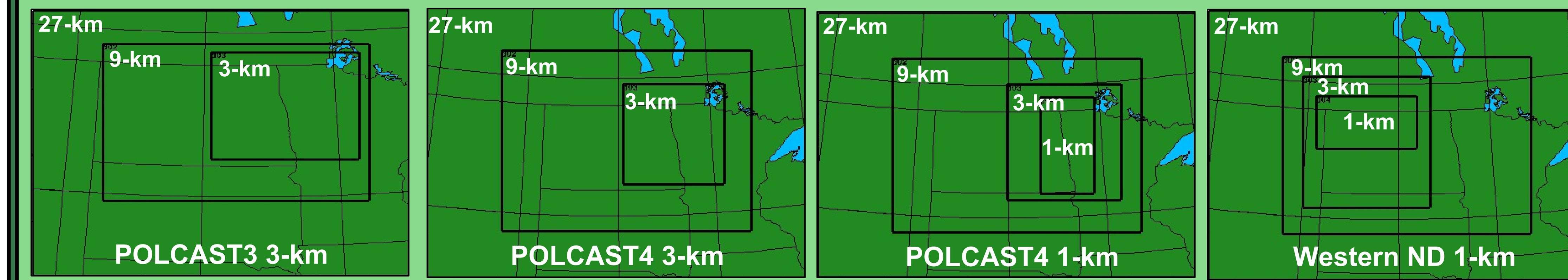
Traditional Skill Scores

A sample forecast and matching observation are seen on the left (generated by MODE). A subjective analysis would generally yield the forecast being very accurate, however according to traditional skill scores (with 1 being a perfect forecast) the forecast is deemed unskillful, as seen in the histogram.

MODE's object merging and matching doesn't always reflect how a person would subjectively group objects for high resolution data There is a need for more development of object-based forecast verification methods for cloud-resolving scales, in order to account for spatial and temporal differences.

Methodology

- **Model data**
  - Simulated reflectivity computed using mixing ratios of rain, snow, and graupel (Stoelinga 2005).
  - Interpolated to 1-km height (AGL) using cubic spline interpolation.
  - For western ND, 3-km data used for the analysis.
  - Eastern ND 1-km data interpolated to a grid spacing of 3-km to have a direct comparison against the other 3-km models.
- **Radar Data**
  - 1-km Constant Altitude Plan Position Indicator (CAPPI) at 2-km horizontal resolution.
  - Interpolated to horizontal 3-km model grid.
- **Method for Object-based Diagnostic Evaluation (MODE) tool, available as part of the Developmental Testbed Center's (DTC) Model Evaluation Tools (MET) package**
  - Simulated reflectivity compared against radar reflectivity hourly for forecast hours 03 to 24 UTC (when radar data existed), using an intensity threshold of > 5 dBZ.
  - The total number of matches (hours) where forecast and radar data existed and was compared are: 275 for eastern ND 2010 3-km, 196 for eastern ND 2012 3-km, 179 for eastern ND 2012 1-km, and 1535 for western ND 2012 3-km.
- **Forecast Evaluation Metrics**
  - Method 1: Basic Convective Forecast Skill
    - Hit, false alarm, and miss rates based on whether convection was present in either domain.
    - Shows a general overview of how well the forecast predicted and timed convection.
  - Method 2: Coverage
    - Only computed when convection exists in both domains.
    - Ratio of area covered by convection divided by total area of domain.
    - Shows how well the forecast predicted convective coverage and to some extent the mode of convection.
  - Method 3: Object Sizes
    - Total number of 'objects' in both domains binned according to their sizes.
    - Forecast area masked to match area observed by radar.
    - Aggregated results show whether the forecasted mode of convection was observed including the amount of convective elements present in the domain.



June 21st to July 23rd, 2010 | June 27th to August 6th, 2012 | July 8th to August 6th, 2012 | June 1st to August 31st, 2012

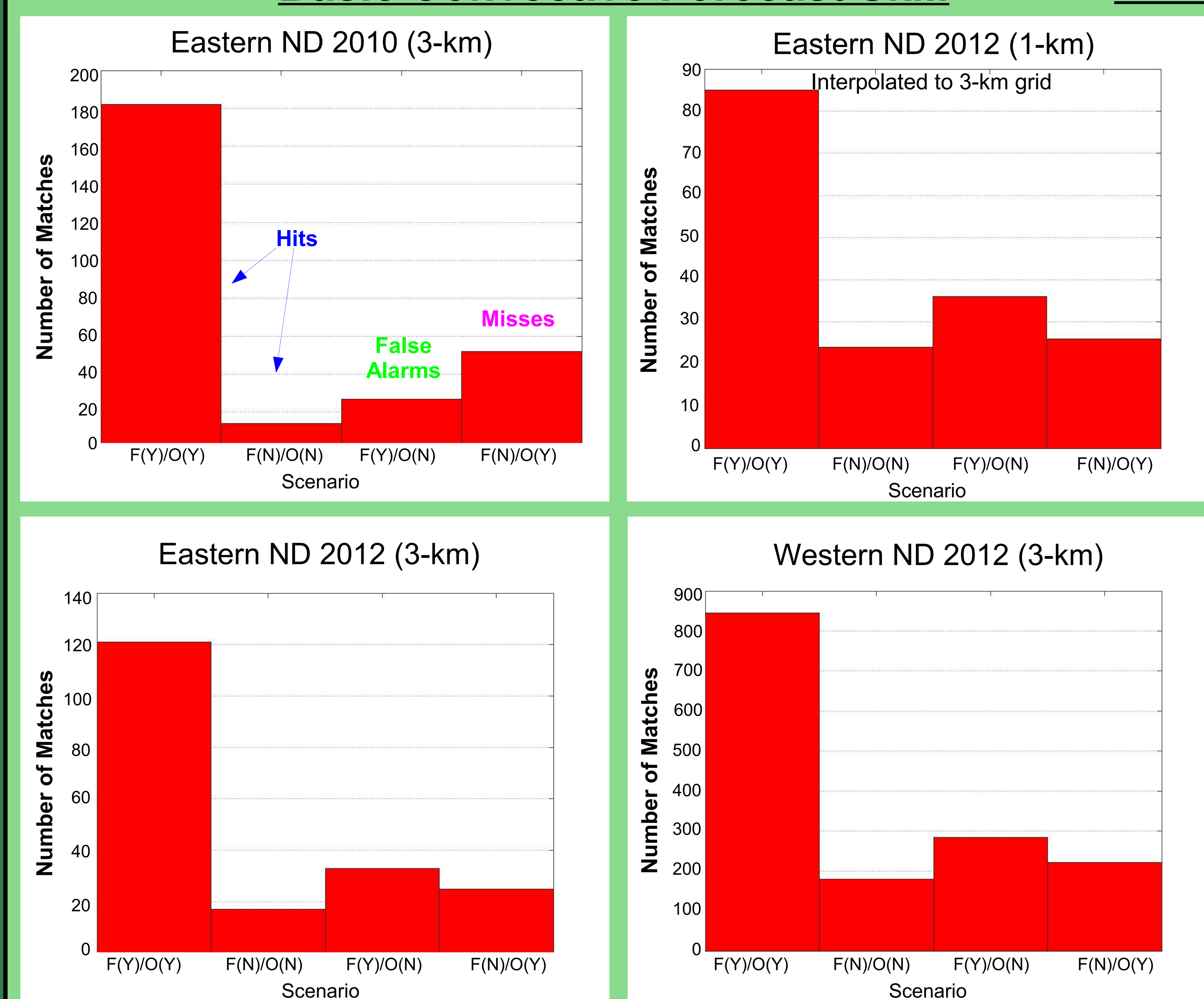
WRF Setup

- WRFv3.1.1
- Initialization:
  - 40-km NAM
  - 00Z
- 44 Vertical levels

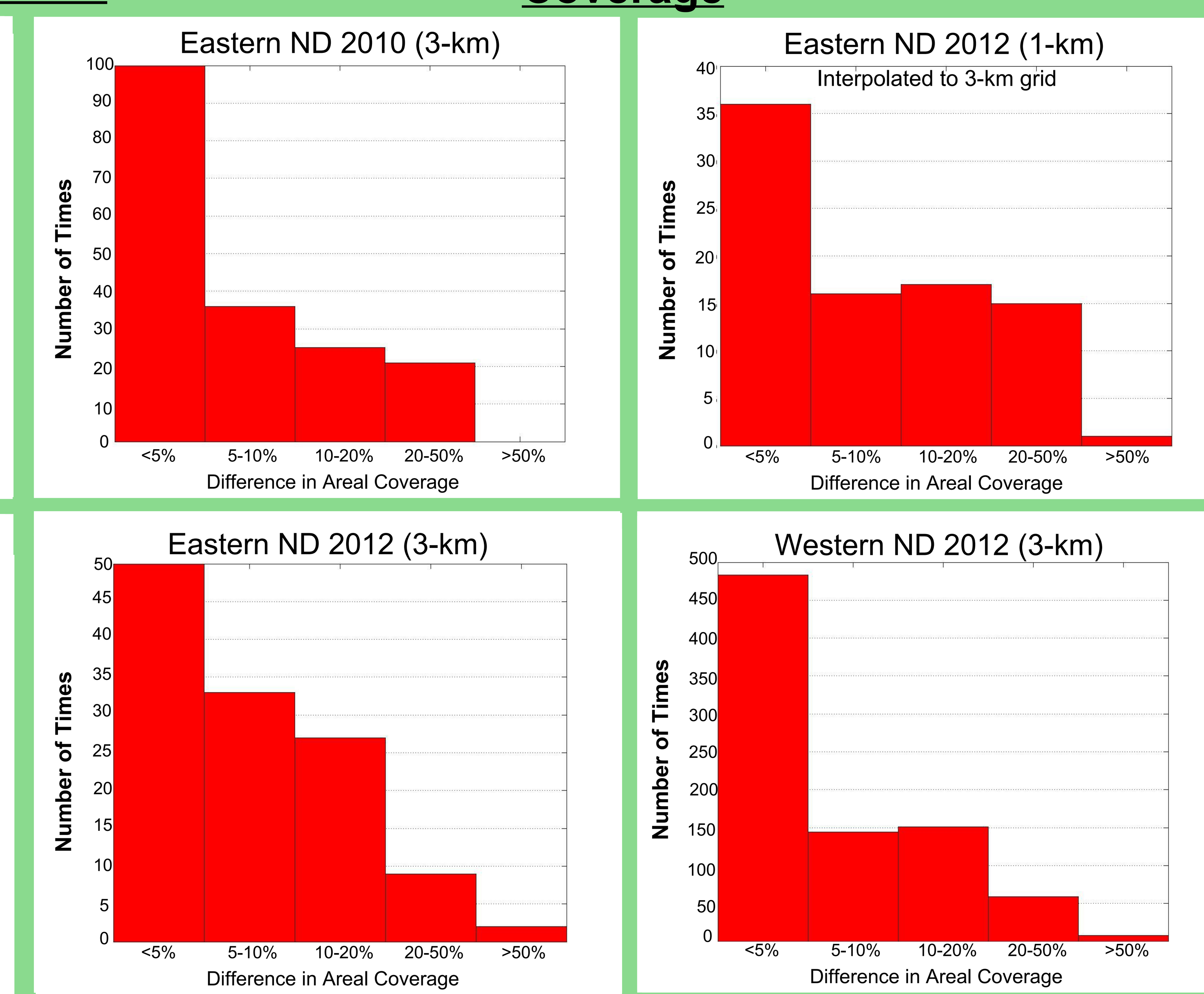
Physics	
Microphysics	WSM6
Longwave	RRTM
Shortwave	Dudhia
Surface Layer	MM5 Similarity
Land Surface	Noah
PBL	YSU
CP	Kain-Fritsch

Results

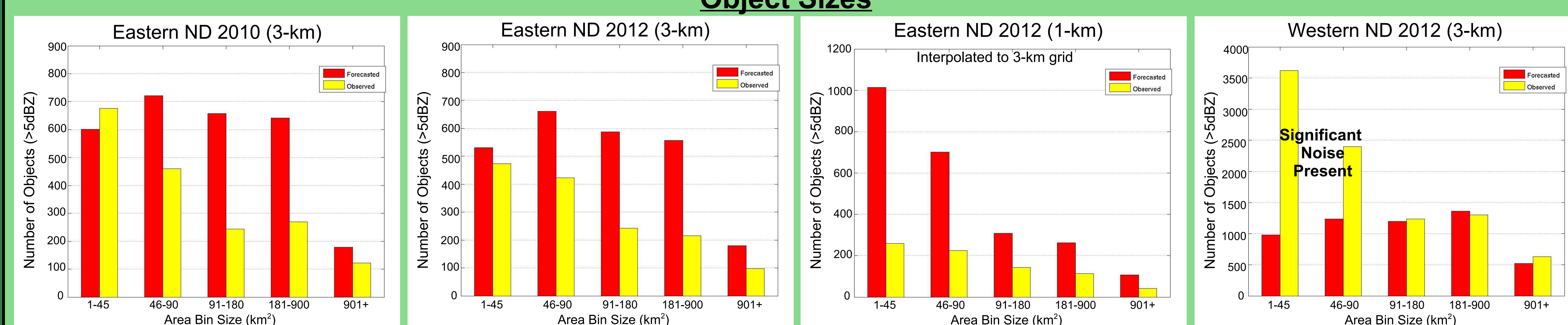
Basic Convective Forecast Skill



Coverage



Object Sizes



- For the 3-km runs, around 70% of cases were hits and, for the subset of hits, the amount of convective coverage forecasted was within 10% of the observed convective area for the majority of the hours analyzed.
- For the 3-km eastern ND runs, objects with sizes of 46 to 900 km^2 were over-forecasted. The 1-km eastern ND runs did slightly better for the larger area bins but significantly over-forecasted smaller objects. Western ND radar data contained significant smaller scale noise, making the smaller bin comparisons unusable. However, western ND runs were the most skillful for predicting large object sizes.
- The 2010 eastern ND forecasts had the highest miss rate and they also under-predicted the smallest objects. All the other simulations had higher false alarm rates, and also over-predicted the smallest object sizes. Since smaller objects are generally surface-based, this may be attributed to the model having trouble correctly forecasting weaker forcing events.
- The 1-km eastern ND runs significantly over-predicted smaller scale convection, but did slightly better with large objects than the 3-km eastern runs. However, the 1-km runs had the highest percentage of area differences in the 20 to 50% bin, showing there were large differences in coverage between domains. The 1-km may have forecasted numerous small cells as opposed to developing fewer but larger features.

Conclusions

- Overall, all the 3-km forecasts did fairly well in forecasting the timing and coverage of convection. Both the 2010 and 2012 eastern ND 3-km runs over-forecasted larger scale convection, while the western ND 3-km forecasted larger features better.
- The 1-km runs were strongly biased towards developing smaller objects.
- Results from eastern ND 3-km simulations suggest that the model may have a hard time dealing with weaker forcing cases and convective initiation.

Future Work

- Evaluate the magnitude of forecasted convection by increasing intensity thresholds in MODE. Comparing the area after applying higher intensity thresholds will show whether the forecast is skillful in predicting the intensity of convection.
- Determine if there is a correlation between forecast skill and cloud concentration nuclei (CCN) concentrations. Many if not all single-moment microphysical schemes assume a constant cloud droplet concentration, which implies a constant CCN concentration. Measurements taken during POLCAST4 show significant CCN concentration variations.
- Currently running several simultaneous WRF realizations with different microphysical packages and different cores. These runs can be used to perform a sensitivity study of the impacts of physics on forecast skill.

Acknowledgments

This project was funded by the North Dakota Atmospheric Resource Board. Special thanks to the National Center for Atmospheric Research for providing computational resources and technical assistance with additional WRF simulations.