Adapting MODE for High-Resolution Model Snowband Verification

Gary Lackmann and Jacob Radford

Department of Marine, Earth, and Atmospheric Sciences

North Carolina State University

Final Project Report to Developmental Testbed Center (DTC) Visitor Program, 2018

Visit: 5-11 June 2018

Report: June 2019

1. Purpose and objective of visit

Mesoscale snow bands in extratropical winter storms produce major societal impacts, and are known to present difficult forecasting challenges (e.g., Novak et al. 2006; Novak and Colle 2012). High-resolution numerical weather prediction (NWP) models such as the high-resolution rapid-refresh (HRRR) model, are able to explicitly represent high-impact mesoscale phenomena such as snow bands. As a result, forecasters are increasingly using this high-resolution model guidance in operational prediction of winter weather hazards such as snow bands. However, to date there have been few studies of the predictability of mesoscale snow bands by models such as the HRRR. An important research goal is to provide operational forecasters with context regarding the capabilities, strengths, weaknesses, and process representation in these models during mesoscale snow bands.

The Model Evaluation Toolkit (MET) software package offers many useful capabilities for a task such as this. Evaluation of model performance for mesoscale features such as snow bands requires the use of object-oriented approaches, such as MODE, which is a component of the MET package. The purpose of this visit to the DTC was to allow the visitor (graduate student Jacob Radford) to collaborate directly with the MET software team to apply MET and MODE to the specific task of snowband verification. Jacob's week-long visit proved to be highly productive and beneficial, as is evident from the use of MET and MODE in his thesis and journal article.

2. Publications

First, this visit was highly beneficial to graduate student Jacob Radford's MS thesis project, which he has since completed. A journal article submitted to *Weather and Forecasting*, is currently undergoing minor revisions. Jacob will continue to utilize MET in his ongoing PhD research.

Radford, J. T., 2019: Verification of high-resolution banded snowfall forecasts. M.S. Thesis, North Carolina State University, available: https://repository.lib.ncsu.edu/handle/1840.20/36330

Radford, J. T., G. M. Lackmann, and M. A. Baxter, 2019: An evaluation of snowband predictability in the High-Resolution Rapid Refresh. *Wea. Forecasting*, in revision (minor revisions).

Acknowledgement in the *Weather and Forecasting* article: "NCAR and the DTC provided MODE software and Tara Jensen, Jamie Wolff, John Gotway, and Randy Bullock provided advice via the Visiting Scientist Program."

Additionally, Jacob Radford provided a summary of this work in the DTC newsletter. The results section below is based in part on the newsletter article, along with results in the *Weather and Forecasting* paper.

3. Results

Jacob Radford, a graduate student at North Carolina State University, visited the DTC during June of 2018, investigating the forecast skill of the High-Resolution Rapid Refresh (HRRR) model for banded snowfall events. In particular, he evaluated the HRRR's ability to capture the location, areal extent, orientation, and aspect ratio of these locally enhanced regions of reflectivity. In theory, snowbands should be adequately resolved by the HRRR thanks to its fine grid-spacing, but model skill has not yet been assessed quantitatively.

Snowbands, or narrow regions of intense snowfall, present hazardous travel conditions due to rapid onset, high precipitation rates, and reduced visibility. Though there has been no quantification of the societal or economic impacts associated with snowbands, in particular, the economic costs of heavy snow events are estimated to be in the billions. Furthermore, mesoscale precipitation bands account for a significant portion of annual precipitation and were found to occur in a majority of cold-season precipitation events in the Northeast and Central U.S. Because these snowbands are small in scale, even the mere occurrence of snowbands is difficult to predict, not to mention the timing, location, and intensity. Thus, there is great incentive to improve understanding of the environmental conditions, physical processes, climatologies, and predictability of snowbands.

Jacob utilized the DTC's Method for Object-based Diagnostic Evaluation (MODE) to match snowbands in the HRRR's 1000-m reflectivity field to bands in national mosaicked base reflectivity fields. However, snowbands were defined based on local reflectivity heterogeneity rather than a set reflectivity threshold, a capability not possible in the current iteration of MODE. Thus, the primary goal of Jacob's visit was to enable greater flexibility in MODE object identification. Jacob worked with Model Evaluation Tools (MET) developers Tara Jenson, Jamie Wolff, John Gotway, and Randy Bullock to implement these changes in MODE. The team quickly determined that the most practical way to accomplish this feat was to build a Python interface for MODE, allowing users to define objects in a Python script however they see fit and then input these objects into MODE via Numpy arrays or xarrays. This added functionality is a significant stepping stone for MODE user flexibility and is available as of MET v8.0.

With MET v8.0, Jacob could then define snowbands as a narrow region of reflectivity at least 1.25 standard deviations above the local reflectivity background, identified in HRRR-simulated and observed reflectivity, and matched with MODE based on similarity in location, area, orientation angle, and aspect ratio. The distributions of interest scores, or measures of the observation/forecast similarities based on these four properties, are shown in Fig. 1. The median interest score of 0.66 indicates that while the HRRR demonstrates some ability to match observed snowband cases, there are significant errors in at least two of the four interest parameters. Applying a cutoff of 0.70 to correspond to a reasonably well-forecasted case, only 30% of cases were well-forecasted by the HRRR. Ultimately, while the HRRR may be helpful in identifying areas of heightened snowband risk, it lacks forecast precision in location and timing. The next step in this work will be to apply a similar verification procedure to the HRRR ensemble to evaluate probabilistic snowband forecast skill.

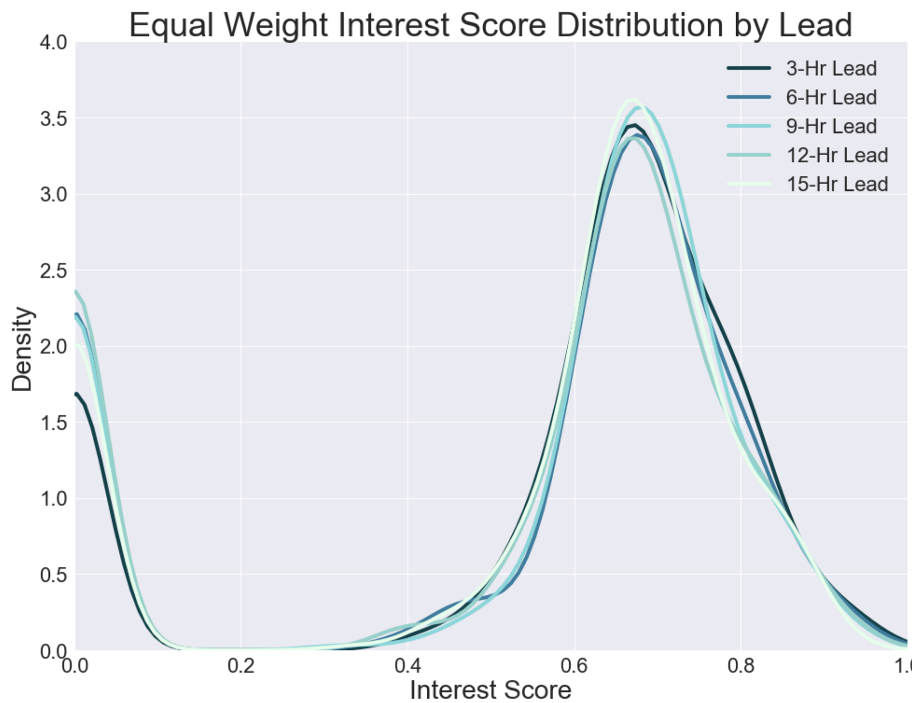

Figure 1: Distributions of mean MODE interest scores at five different forecast leads (varying shades of blue).

As mentioned above, Jacob's work required an object-oriented approach to model verification, and MODE allowed us to accomplish this, aided in large part by Jacob's visit to DTC. For additional details of research results from the HRRR verification study, please see the upcoming Weather and Forecasting paper. Here is an excerpt from that paper: "Fuzzy verification does not provide information on whether matched bands exhibit similar properties. The Developmental Testbed Center's (DTC) Method for Object-based Diagnostic Evaluation (MODE) tool lets researchers match precipitation features based on the similarity between forecast and observed object properties (Davis et al. 2009). The choice of these parameters and their relative importance is somewhat arbitrary, but should be tailored to the specific phenomenon in question. The parameters we deemed most important to band forecast skill were centroid distance, area ratio, aspect ratio difference, and orientation angle difference, which we weighted equally. A tunable interest function was developed to produce an overall interest score (a dimensionless measure of forecast and observed object similarities) based on the four parameters[1].

$$I_{of} = C_D \frac{150.0 - D_{of}}{600.0} + \frac{2.0 - AS_{of}}{8.0} + C_O \frac{90° - O_{of}}{360°} + \frac{1.0 - A_{of}}{4.0} \quad (1)$$

$$C_D = Centroid\ Distance\ Confidence\ ;\ D_{of} = Centroid\ Distance$$

$$AS_{of} = Aspect\ Ratio\ Difference\ ;\ C_O = Orientation\ Angle\ Difference\ Confidence$$

[1]We refer the reader to the Model Evaluation Tools user documentation (Gotway et al. 2018) for more details on how MODE calculates interest scores.

$$O_{of} = Orientation\ Angle\ Difference\ ;\ A_{of} = Area\ Ratio\ Difference$$

This interest function was designed specifically so that an interest score of 0.70 or greater correlates to a reasonably well-forecast banding event, similar to methods employed by Davis et al. 2009. If errors were distributed equally between the four interest parameters, the 0.70 cutoff would equate to a centroid distance of 45 km, an area ratio of 0.70, an aspect ratio difference of 0.60, and an orientation angle difference of 27°. The mean interest score for all 357 cases across all lead times was 0.54, while the median interest score was 0.66. The distribution [Fig. 2 (15 in paper)] is highly skewed towards lower values due to the prevalence of cases in which one or more of the leads forecast no intense precipitation objects, leading to interest scores of 0.0. Breaking down the distribution by lead [Fig. 1 (16 in paper)], there is no trend with decreasing lead times, consistent with what was found in the traditional and fuzzy verification sections.

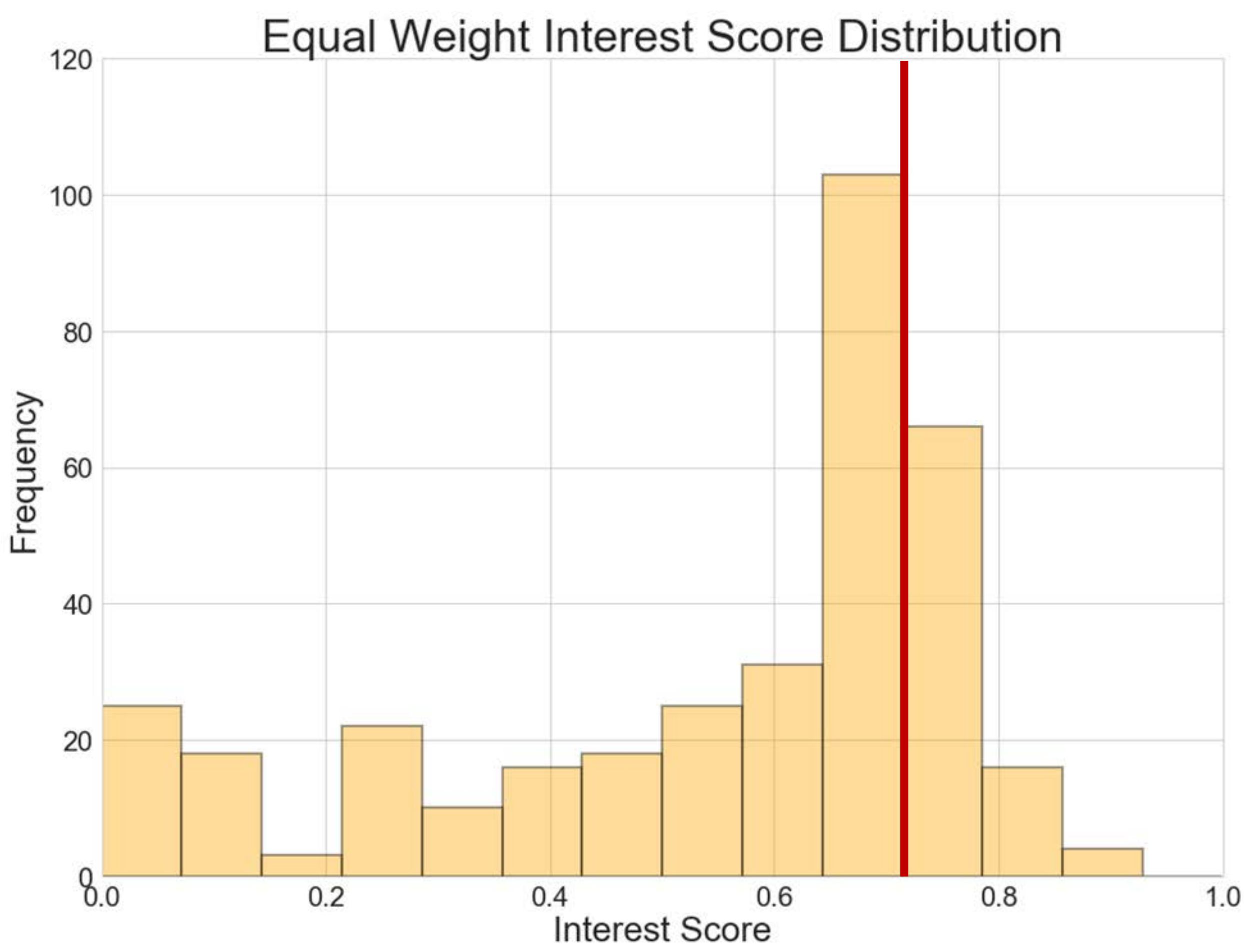

Figure 2: The distribution of (across-lead) mean MODE interest scores between forecast precipitation objects and observed bands using equal weightings of interest parameters. The threshold for a well-forecast case is indicated by the bold red line.

Using the 0.70 “well-forecast” threshold, the mean interest score across leads indicate that 107 out of the 357 cases (30%) were well-forecast. 247 out of the 357 cases (69%) had at least one forecast lead time with an interest greater than 0.70. This is similar to what was found in the traditional verification, in which 66% of cases were forecast in at least one forecast lead. Though less than ideal, this could help to signal forecasters that conditions are favorable for band development. Since there is no consistent, particular, lead time which exhibits improved band forecast performance, it would be difficult for forecasters to determine which forecast lead will be correct; however, use of probabilistic visualizations like a time-lagged ensemble could indicate areas where there is enhanced risk of banding."

Jacob found collaboration with the DTC to be an extremely valuable experience vital to the completion of his Master’s research and a step towards his Ph.D. Everyone at the DTC was extremely friendly, accommodating, and knowledgeable about forecast verification.

Acknowledgements. Another important point of emphasis regards professional development. The project benefit to Jacob Radford extended well beyond accelerated use of MET and MODE, but it gave Jacob a chance to experience a different professional environment (at NCAR), and to network and meet scientists at the DTC. We are grateful for this aspect of the visit as well as the scientific benefits that resulted from it, and the hospitality extended by the DTC scientists made a very positive impression on Jacob.

4. References

Davis, C.A., B.G. Brown, R. Bullock, 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.

Gotway, J.H., Newman, K., Jensen, T., Brown, B., Bullock, R., Fowler, T., 2018: Model evaluation tools version 8.0 (METv8.0) user’s guide, 430 pp.

Novak, D. R., J. S. Waldstreicher, D. Keyser, and L. F. Bosart, 2006: A forecast strategy for anticipating cold season mesoscale band formation within eastern U.S. cyclones. *Wea. Forecasting*, **21**, 3–23.

Novak, D. R., and B. A. Colle, 2012: Diagnosing snowband predictability using a multimodel ensemble system. *Wea. Forecasting*, **27**, 565–585.

Radford, J. T., 2019: Verification of high-resolution banded snowfall forecasts. M.S. Thesis, North Carolina State University, available: https://repository.lib.ncsu.edu/handle/1840.20/36330

Radford, J. T., G. M. Lackmann, and M. A. Baxter, 2019: An evaluation of snowband predictability in the High-Resolution Rapid Refresh. *Wea. Forecasting*, in revision (minor revisions).