

# **Object-based Verification for the HRRR Model Using Simulated and Observed GOES Infrared Brightness Temperatures**

Jason A. Otkin, Sarah Griffin, and Chris Rozoff

Cooperative Institute for Meteorological Satellite Studies

Space Science and Engineering Center

University of Wisconsin-Madison

Email: [jason.otkin@ssec.wisc.edu](mailto:jason.otkin@ssec.wisc.edu)

Final Project Report

01 July 2016

## 1. Introduction

Satellite infrared brightness temperatures (BT) provide detailed information about the horizontal and vertical distribution of clouds and have therefore been used in prior studies to evaluate the accuracy of the cloud field in high-resolution numerical weather prediction model forecasts (e.g. Otkin and Greenwald 2008; Otkin et al. 2009; Cintineo et al. 2014; Thompson et al. 2016). Given the fine spatial resolution of the High Resolution Rapid Refresh (HRRR) model, qualitatively small differences between the observed and forecast cloud fields may lead to large quantitative errors that differ greatly from a subjective forecaster evaluation of the model forecast accuracy. These differences can be quantified using traditional (e.g. root mean square error, bias, etc.) or neighborhood-based (e.g. fractions skill score) metrics. However, though these verification methods provide useful information concerning the model accuracy, additional information about errors in the spatial distribution of the cloud field can be obtained through the use of more sophisticated object-based verification tools such as the Method for Object-Based Diagnostic Evaluation (MODE; Davis et al. 2006a, b).

Few studies have used satellite observations and object-based verification tools such as MODE to evaluate the accuracy of the simulated cloud field. An exception is a recent study by Mittermaier and Bullock (2013) that used the MODE system to assess the spatial and temporal characteristics of cloud cover forecasts in high-resolution models over the United Kingdom. They showed that object-based verification provides a useful way to evaluate different attributes of the forecast cloud field. For example, object-based verification can provide information about the simulated cloud structure, such as errors in the location, shape, orientation, and spatial extent of the cloud field that is not readily obtained using traditional or neighborhood-based verification metrics. Therefore, object-based statistics provide a more advanced assessment of the forecast

accuracy. The aim of this project was to learn the intricacies of MODE and to investigate how MODE output could be used to analyze the accuracy of the experimental HRRR (HRRRx) model through comparison of observed and simulated cloud objects.

## 2. Data

### a. *High Resolution Rapid Refresh model simulated brightness temperatures*

This project used simulated Geostationary Observing Environmental Satellite (GOES) 10.7  $\mu\text{m}$  BTs computed using output from the HRRRx model. The 10.7  $\mu\text{m}$  BTs are sensitive to cloud top properties when clouds are present and to surface skin temperature when clouds are absent from a given scene. The HRRRx is an hourly-updating model that covers the contiguous United States (CONUS) with 3-km horizontal grid spacing and 51 vertical levels. HRRRx simulated BTs were available for forecast hours (FHs) 0 to 24 during the study period. The version of the HRRRx model used for this study was implemented at Earth System Research Laboratory on 4 May 2015 (Earth System Research Laboratory 2016). It uses initial conditions from the Rapid Refresh (RAP) model and then applies data assimilation at 3-km resolution including the assimilation of radar reflectivity. The HRRRx is a convection-allowing model that does not include deep convective parameterization (Benjamin et al. 2016). The HRRRx uses the Thompson-aerosol aware v3.6.1 microphysics scheme, MYNN v3.6+ planetary boundary layer scheme, RUC land surface model, and RRTMG shortwave and longwave radiation schemes (Earth System Research Laboratory 2016).

Simulated GOES BTs were computed for each forecast time using HRRRx model output and the Community Radiative Transfer Model (CRTM, Han et al. 2006) in the Unified Post Processor (UPP). For clear grid points, simulated BTs are computed using several model-predicted fields, such as surface skin temperature, 10-m wind speed, and vertical profiles of temperature and water vapor. For cloudy grid points, additional information about cloud radiative properties is required to calculate simulated BTs. Vertical profiles of mixing ratio and number concentration are used to estimate the properties for each cloud hydrometeor species (cloud water, cloud ice, rain water, snow and graupel) predicted by the Thompson-aerosol aware microphysics scheme. In the CRTM, standard look-up tables for cloud absorption and scattering properties, such as extinction, single-scatter albedo, and the full scattering phase function are used to assign values to each hydrometeor species as a function of the cloud effective diameter computed using the particle size distribution assumptions for that scheme (e.g. Otkin et al. 2007). Cloud optical properties are then computed for each species and model layer in the vertical profile before computing the simulated infrared BTs.

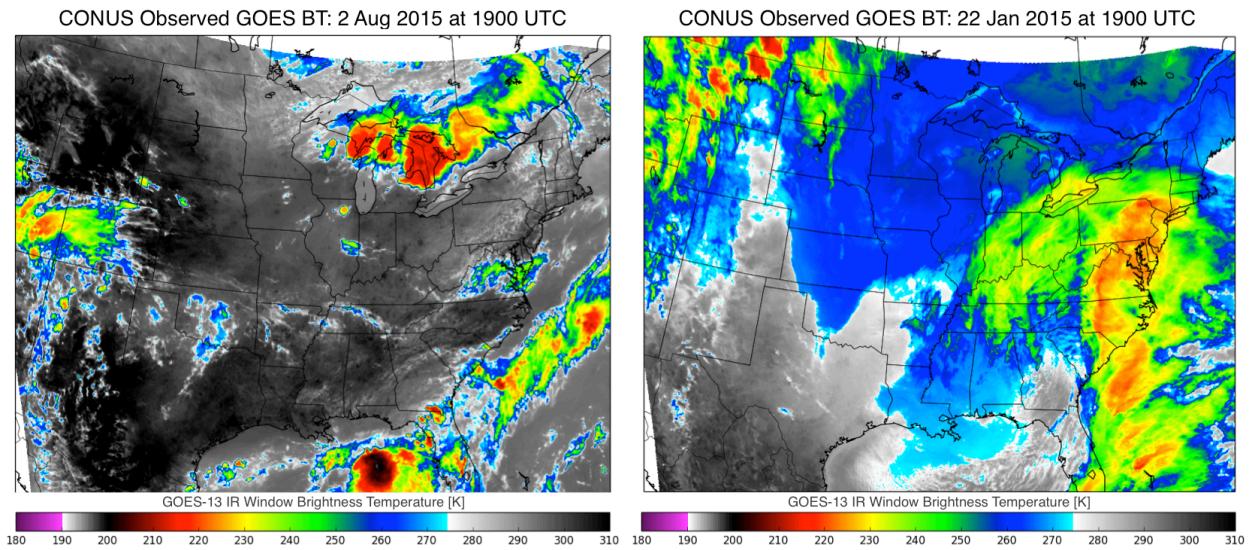
*b. Observed brightness temperatures*

The satellite validation data used during this study is from the GOES 13 imager. The 10.7  $\mu\text{m}$  GOES BTs have a 4-km spatial resolution at nadir and are remapped to the 3-km HRRRx grid using a weighted average of all the observed pixels overlapping a given HRRRx model grid box. The GOES imager typically completes a scan over CONUS every 15 minutes except for every 3 hours (00, 03 UTC, etc.) when the scan at the top of the hour is skipped in order to complete a full disk scan. Thus, simulated HRRR BTs will be compared to the 0 minute

scan for most hours, but will be compared to the scan starting 15 minutes prior to the HRRRx forecast time for cases when the 0 minute CONUS scan is skipped. This occurs at 3-h intervals. While this may introduce some uncertainty in the analysis, it is expected to be minor.

### *c. Seasonal comparison*

To assess the HRRRx forecast accuracy, simulated and observed GOES 10.7  $\mu\text{m}$  BTs from two one-month long periods, including 01-31 August 2015 and 01-31 January 2016, were evaluated. These time periods were chosen so that the HRRRx model forecast accuracy could be assessed during both warm and cool seasons given potential differences in cloud characteristics. Representative snapshots of the GOES 10.7  $\mu\text{m}$  BT are shown in Figure 1. The image on the left is from 02 August 2015, while the image on the right is from 22 January 2016. Cloud objects in the upper troposphere associated with the coldest BTs were generally smaller in August than in January, though both large and small objects occurred during both months.



*Figure 1. Observed GOES 10.7  $\mu\text{m}$  BT images valid at 1900 UTC on 02 August 2015 (left image) and 1900 UTC on 22 January 2016 (right image).*

### **3. Method for Object-Based Diagnostic Evaluation**

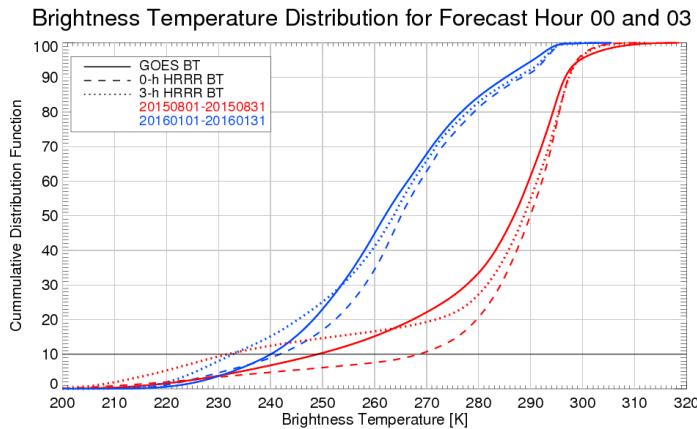
MODE is used to analyze the accuracy of the HRRRx forecast cloud field. MODE is a technique for identifying and matching objects in two different fields (Davis et al. 2006a, b, 2009). Objects are meant to represent “regions of interest”, which for this project were chosen to be upper-level cloud features containing the coldest infrared BTs. The MODE process is fully described in Davis et al. (2006a); however, a short outline of the process as applied to the infrared BT datasets is provided here for context:

1. Spatially smooth the forecast and observed BT fields using a process called convolution thresholding to identify objects.
2. Calculate various object attributes for each observed and forecast cloud object.
3. Match the forecast and observed cloud objects using a fuzzy logic algorithm and then calculate various attributes of the paired objects, such as intersection area and distance.
4. Write out attributes for individual objects and matched object pairs for assessment.

The convolution radius used for both the observed and forecast fields in step one was chosen to be 5 grid points. This allows for the analysis of small-scale storms, as a range of 2 to 8 grid points is stated by Cai and Dumais (2015) as identifying convective storm objects in ~4-km resolution radar imagery. Object merging in the individual observed and forecast fields was not performed during this study.

Step one of the above process also requires choosing a BT threshold to define the edges of the cloud objects. However, as seen in Fig. 1, cloud objects can contain different BTs depending upon the season, weather regime, and location. In addition, Griffin et al. (2016) identified a cold bias in the HRRRx forecast BTs that varies as a function of forecast hour. To

illustrate these differences, probability distributions for the observed and simulated 10.7  $\mu\text{m}$  BT from August 2015 and January 2016 at forecast hours 0 and 3 are shown in Fig. 2. Although the coldest BTs occurred during August, the entire BT distribution is overall colder in January than it was during August. In addition, the coldest HRRRx simulated BTs were colder in the 3-h forecast than in the 0-h forecast and moreover the sign of the bias for BTs  $< 260$  K changes from the initialization to the 3-h forecast. Thus, to account for these seasonal and forecast hour differences, the 10<sup>th</sup> percentile of the BT distributions was chosen to define the cloud edges instead of an arbitrary BT threshold. The 10<sup>th</sup> percentile is used so that the analysis focuses on the coldest cloud objects occurring in the upper troposphere. Ten days, including the current time, are used to determine the 10<sup>th</sup> percentile BT threshold for upper level clouds because it is



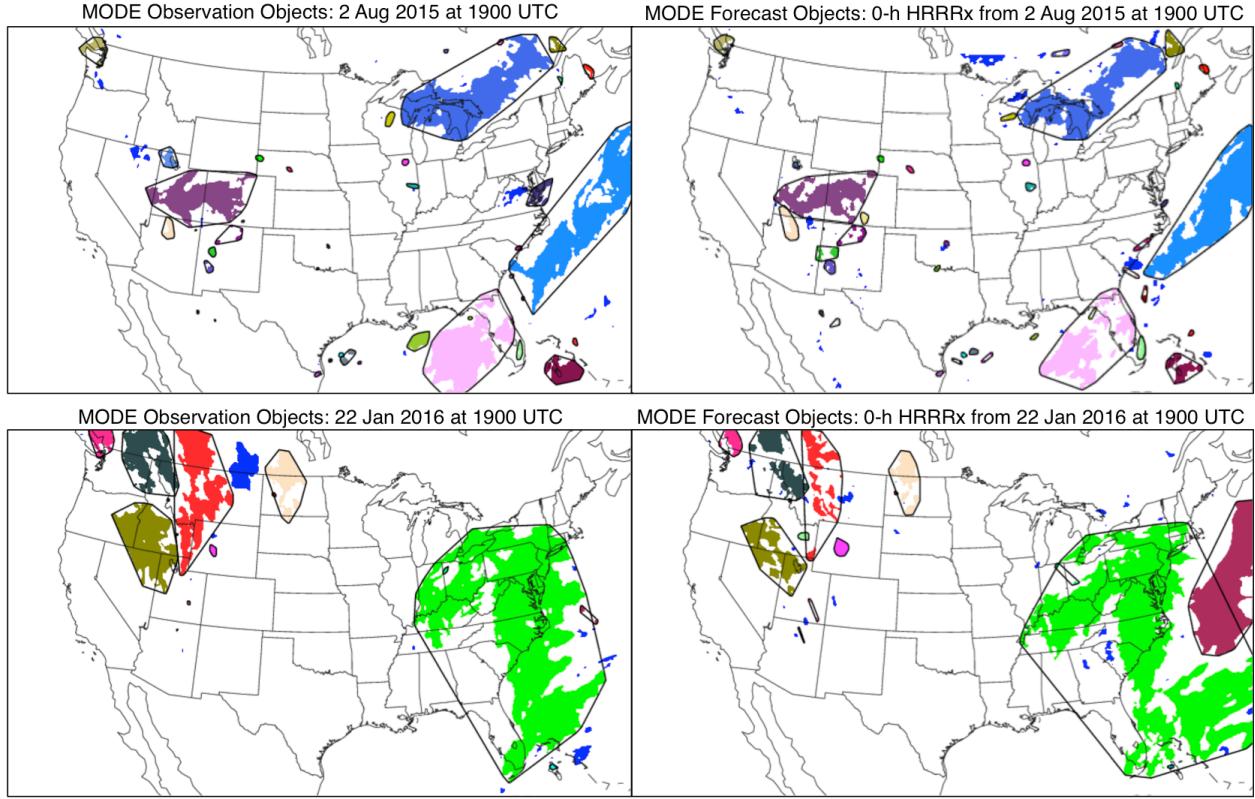
*Figure 2. Observed (solid) and forecast GOES 10.7- $\mu\text{m}$  BT probability distributions for August 2015 (red) and January 2015 (blue) for the 0-hr HRRRx initialization (dashed) and the 3-hr HRRRx forecast (dotted).*

account for any potential variations between different HRRRx initialization times.

Sample MODE objects defined using the observed and simulated BT thresholds from the 0-h HRRRx forecasts are shown in Figure 3. For each object pair, MODE computes an interest

possible that no upper level clouds are present if a shorter time period is used. The observed BT threshold is calculated using observations valid at the same time of day to account for differences in the cloud characteristics during the diurnal cycle. The forecast BT threshold is calculated using the HRRRx simulated BT from the same initialization time and forecast hour to

value portraying the correspondence between the two objects. The interest value is a weighted combination of the object pair attributes. Interest values range from 0 to 1, with a perfect match having an interest value of 1. When calculating the interest value, MODE uses various object pair attributes whose weights are defined by the user. The attributes and user-defined weights used for this project are shown in Table 1. Overall, this analysis prioritizes the distance and size comparison between the objects. Two distance attributes are highly weighted relative to other attributes, with the minimum distance (`boundary_dist`) between objects having a slightly lower weight than the centroid distance (`centroid_dist`). This was done to put more emphasis on the displacement between the objects' centers of mass rather than object overlapping because large objects can have a sizeable center displacement but still overlap. However, it should be noted that because the MODE centroid distance weight is the user-defined weight multiplied by the ratio of the objects' areas, the boundary distance between objects has greater weight when the ratio between the observed and forecast area is less than 0.75. The area attributes receive the same user-defined weight as the distance attributes. The ratio of the objects' areas (`area_ratio`) is higher than the ratio of the intersection area of objects to the observation/forecast object's area (`int_area_ratio`) because the `int_area_ratio` value can be artificially high when a small object is fully enclosed within a larger object.



*Figure 3. MODE identified observed cloud objects (left panels) and HRRR-x 0-h analysis cloud objects (right panels) valid at 1900 UTC on 02 August 2015 (top row) and at 1900 UTC on 22 January 2016 (bottom row).*

Object Pair Attribute	User-Defined Weight (%)	Description
centroid_dist	4 (25.0)	Distance between objects' "centers of mass"
boundary_dist	3 (18.75)	Minimum distance between the objects
convex_hull_dist	1 (6.25)	Minimum distance between the polygons surrounding the objects
angle_diff	1 (6.25)	Orientation angle difference
area_ratio	4 (25.0)	Ratio of the forecast and observed object areas (whichever yields a lower value)
int_area_ratio	3 (18.75)	Ratio of observed (forecast) object to the objects' intersection area (whichever yields a higher value)

*Table 1. User-defined MODE object attribute weights used to compute the MODE interest score and brief a description of the object pair attributes used in this analysis.*

#### 4. Mode Skill Score

Because MODE generates extensive output, we explored methods to condense its output into a single value that still preserves unique information provided via object-based verification. To do this, we developed a new verification metric named the MODE Skill Score (MSS). The MSS is defined in Griffin et al. (2016) as an area-weighted calculation using the interest values computed by MODE. It includes two components, one for cloud clusters and the other for cloud objects. It is computed as follows:

$$\text{MSS} = \sum_{i=1}^C \frac{\text{Area}_{\text{Observed Cluster}(i)}}{\text{Total Area}} * \text{Interest Value}(i) + \sum_{j=1}^O \frac{\text{Area}_{\text{Observed Object}(j)}}{\text{Total Area}} * \text{Interest Value}(j) \quad (1)$$

The Total Area in the denominator is defined as the area of all observed cloud objects plus the area of all forecast cloud objects that are unmatched to the observed cloud objects. A cluster is defined as any set of one or more objects in one field that matches any one or more objects in the other field (Developmental Testbed Center 2014). For this project, object matches are defined as clusters if the interest value exceeds 0.65. Individual objects can be components of the same cluster if two or more observed (forecast) cloud objects match the same forecast (observed) object. Because the maximum centroid distance in this analysis is 200 km, this means that any object in the observed field will be matched, and therefore have a non-zero interest value, with a forecast object that is within 200 km. The MSS has a range from 0 to 1, with 1 representing a perfect forecast. A MSS equal to 0 indicates that there were no forecast cloud objects with a centroid distance within 200 km of every observed cloud objects; however, this is unlikely to occur.

The MSS calculation first considers clusters to account for multiple objects in the observed (forecast) field that may correspond to a single object in the forecast (observed) field.

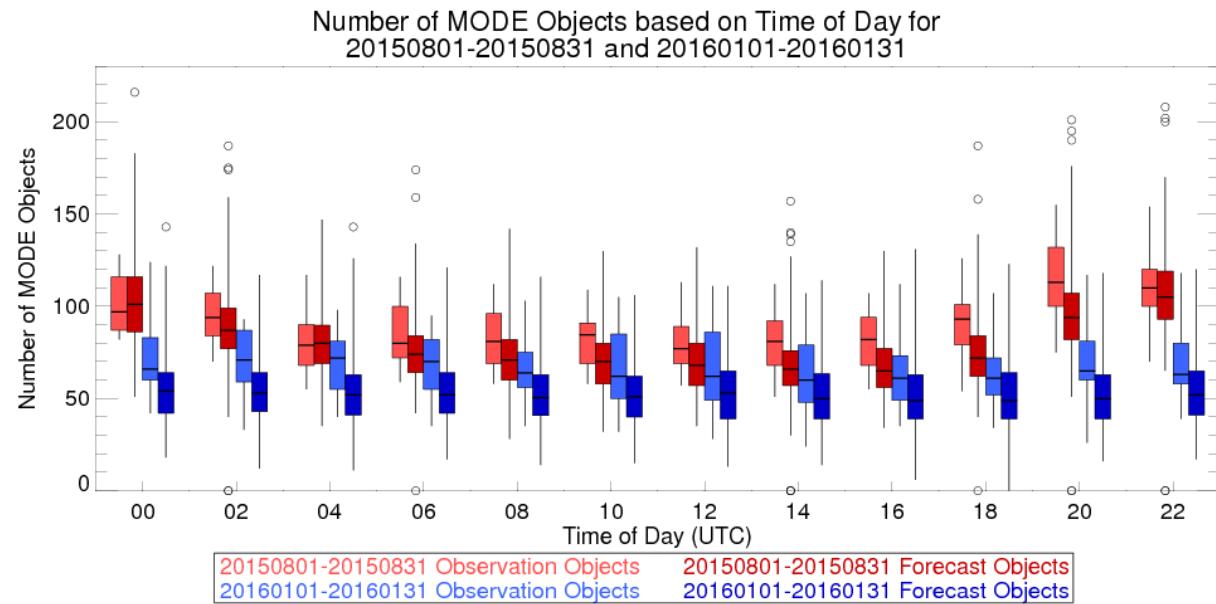
Once the MSS accounts for the clusters, only matching object pairs where both the observed and forecast objects are not already considered in a matched cluster will be used in the MSS calculation. Because each observed and forecast object can only be used once, object matches are analyzed from the highest interest value to the lowest to calculate the highest possible MSS. Object matches with an area ratio less than 5% are not included in the MSS calculation so that large forecast objects are not matched to much smaller observed objects and vice versa.

## 5. Results

### a. *Object number comparison*

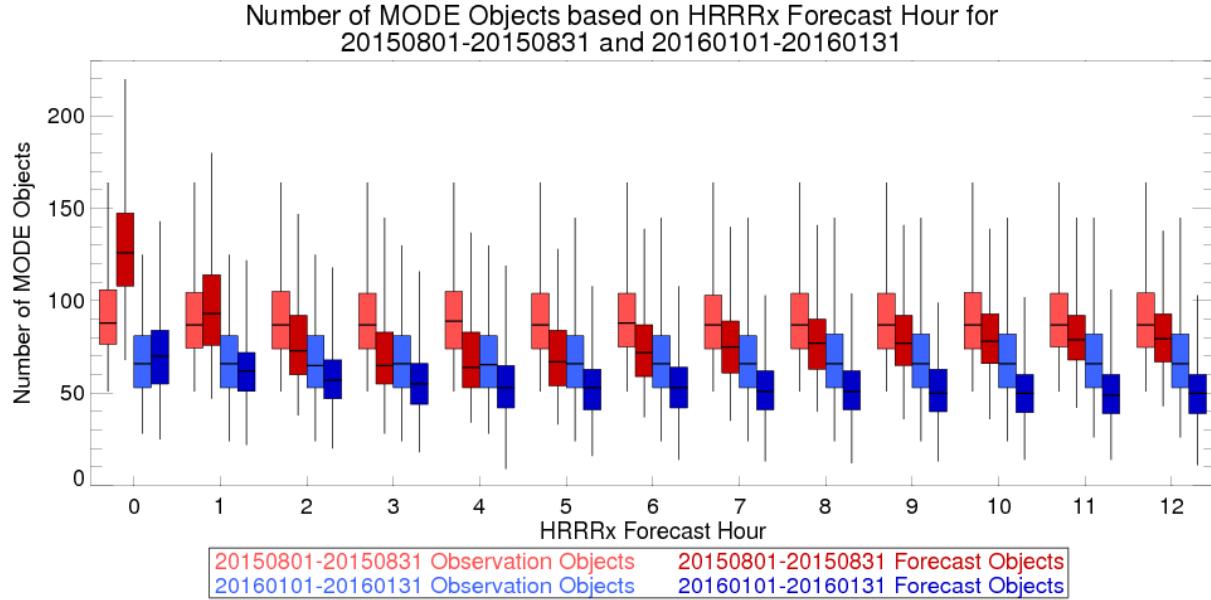
The first step in assessing the accuracy of the HRRRx model forecasts is to determine if it accurately represents the number of upper level cloud objects present in the GOES observations. Figure 4 shows the number of individual observed and forecast objects plotted as a function of time of day using all data from August 2015 and January 2016. Overall, there are more cloud objects in August than in January. Because the total cloud object area in both datasets is equal to 10% of the grid because we are using the 10<sup>th</sup> percentile of the BT distribution to define the edges of the cloud objects, this means that the median cloud object size is smaller during August. The diurnal cycle is also more prominent in August compared to January. Both of these characteristics are consistent with the more predominant small-scale convective cloud features found during the summer compared to the larger synoptic-scale cloud systems more frequently observed during the winter.

Though the different characteristics of the diurnal cycle are accurately captured in both the August and January forecasts, there are typically not enough HRRRx forecast objects in both months. For example, the median number of HRRRx forecast objects is smaller compared to the observed objects for each time of day and season in Fig. 4, with exceptions occurring near 00 and 04 UTC during August. However, even though the median number of forecast objects is smaller, it can also be seen in the box whiskers that the largest number of HRRRx forecast objects is typically larger than the largest number of observed objects. As can be seen in Fig. 5, which shows the number of objects plotted as a function of forecast hour, these occurrences are mostly associated with the 0-h HRRRx initializations when there are typically too many cloud objects. This is due to there being multiple small convective cores in the model analyses versus larger clouds in the observations. In later forecast hours there are generally fewer HRRRx forecast objects than observed in the GOES imagery. The median number of forecast objects



*Figure 4. Box plot diagram depicting the range of the number of observed (lighter colors) and forecast (darker colors) upper level cloud objects identified by MODE for August 2015 (red) and January 2016 (blue) plotted as a function of time of day. Data from all forecast cycles and hours were used to produce the box plot diagram.*

reaches a minimum around FH 4 during August before then slowly increasing with time. During January, however, there is a steady drift toward fewer HRRRx forecast objects at longer forecast lead times.



*Figure 5. Box plot diagram depicting the range of the number of observed (lighter colors) and forecast (darker colors) upper level cloud objects identified by MODE for August 2015 (red) and January 2016 (blue) plotted as a function of forecast hour. Data from all forecast cycles were used to produce the box plot diagram.*

#### b. MODE Skill Score

The HRRRx forecast accuracy was also assessed using the MSS. Because the MSS is computed using the MODE interest scores, and our study assigns equal weight to the object displacement and area attributes, the MSS can be used to assess the accuracy of these object characteristics. Figure 7 shows the mean MSS plotted as a function of forecast hour for August 2015 and January 2016, with the grey indicating the 95% confidence interval around the mean. Overall, the 1-h forecast is the most accurate for both months, with a steady decrease thereafter

as predictability decreases. While forecast hours 2 and beyond have similar accuracy for each season, the 0-h and 1-h HRRRx forecasts are more accurate for August.

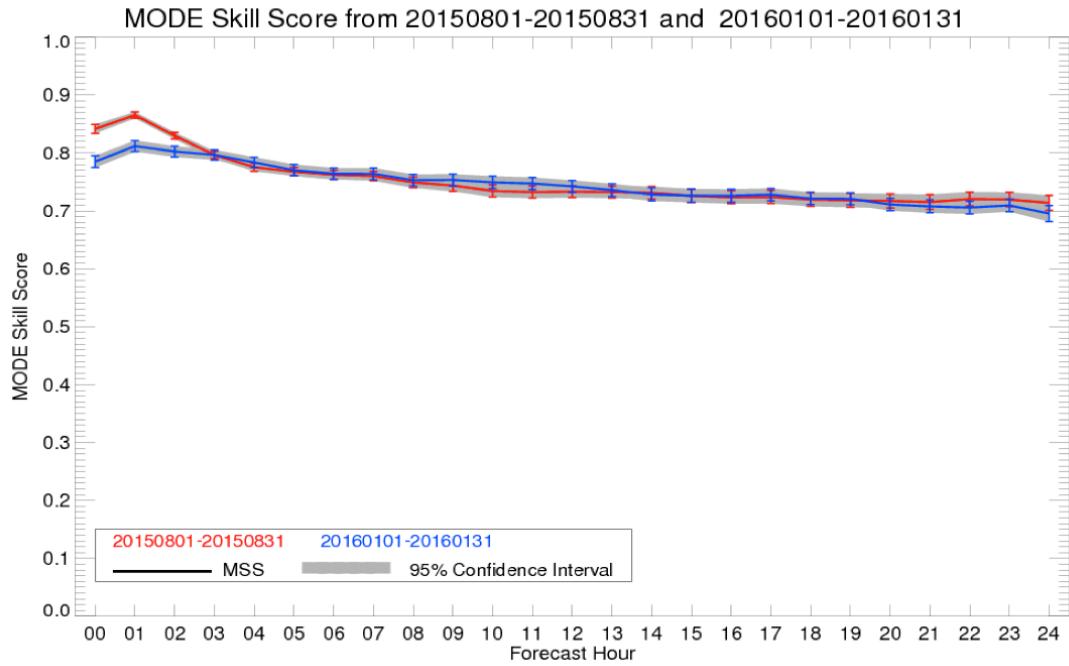


Figure 6. Mean MODE Skill Score values plotted as a function of forecast hour for August 2015 (red line) and January 2016 (blue line).

### c. Why is the 1-h forecast the most accurate?

As shown by the MSS, the 1-h HRRRx forecast is the most accurate forecast for both August and January. To identify why the 1-h forecast is more accurate than the 0-h analysis, the MODE interest scores for the distance and area attributes from the 1-h HRRRx forecasts were compared to those from the 0-h HRRRx initialization. The attribute interest scores for matched objects were arranged based on the size of the matching observed object. To account for the difference in the number of matches based on observed cloud object size, which can be caused

when the same observed object only has a match in either the 0-h or 1-h forecast, the percent of occurrences is calculated. The difference in percent occurrences is then defined as:

$$\text{Difference in Percent of Occurrences} = \frac{\text{Number(score, obs\_size)}_{\text{FH 00}}}{\text{Total(obs\_size)}_{\text{FH 00}}} -$$

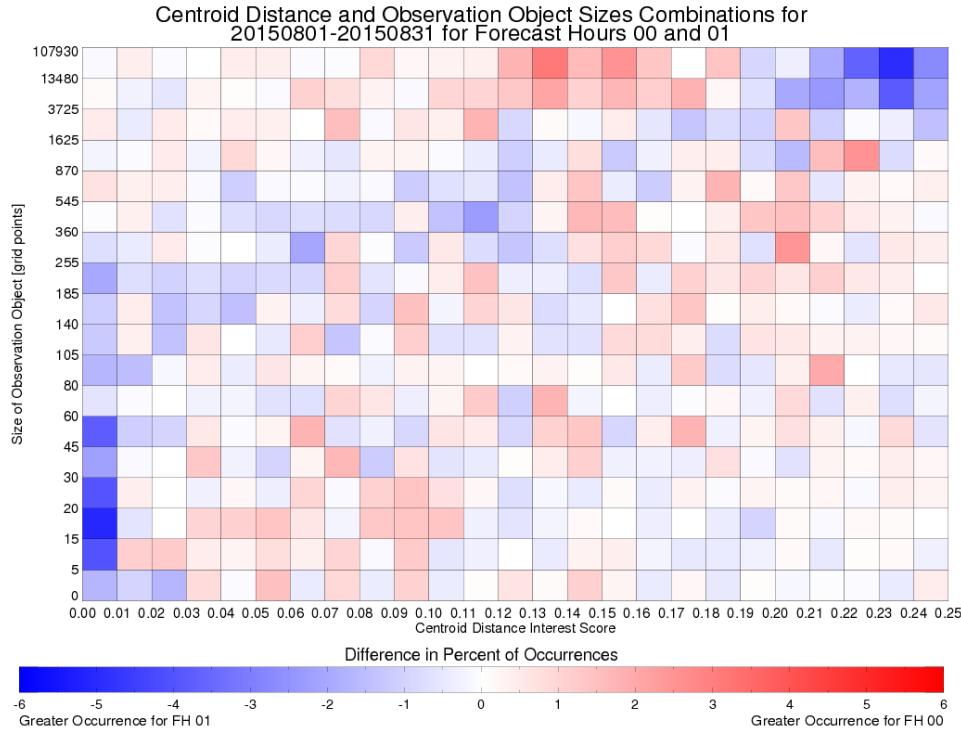
$$\frac{\text{Number(score, obs\_size)}_{\text{FH 01}}}{\text{Total(obs\_size)}_{\text{FH 01}}}$$

where `obs_size` is the size of the observation object in the matched object pair and “score” is the attribute’s interest score.

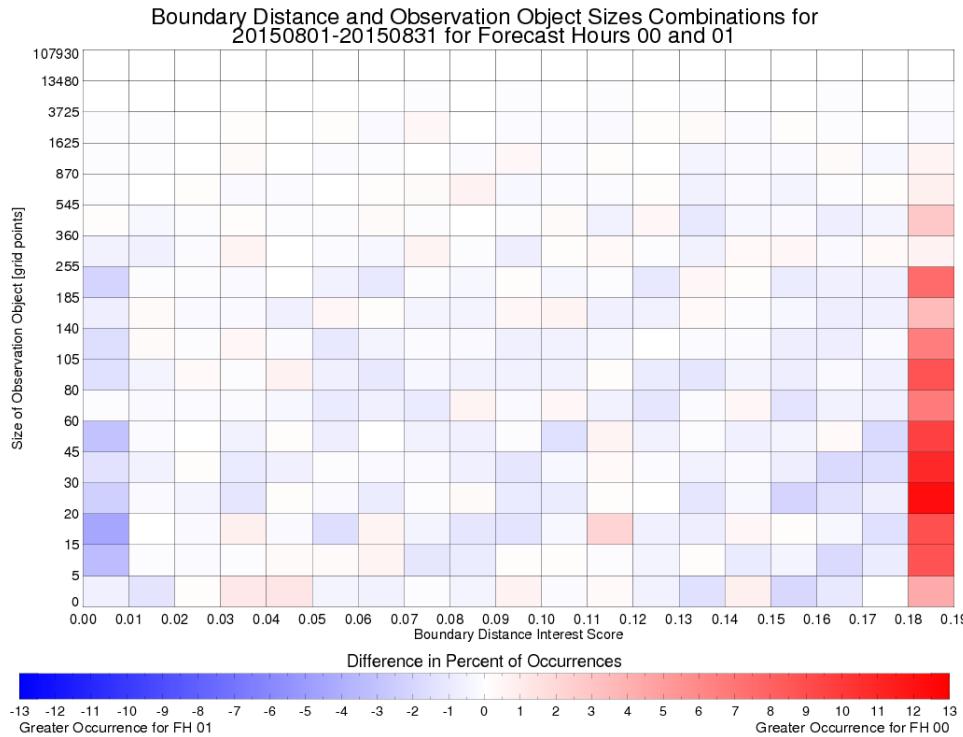
The difference in percent occurrences for the centroid distance attribute is shown in Fig. 7. Blue (red) colors indicate that the centroid distance interest score and observed object size combination occurs more frequently in the FH1 (FH0) object pairs. For small objects, there are more with lower interest scores in the FH1 forecasts than in the FH0 initialization. This is not unexpected because displacement errors tend to increase with increasing forecast hour (Griffin et al. 2016) and lower interest values are associated with greater displacement errors between the objects’ “centers of mass”. However, higher interest scores are also observed more frequently in the FH1 forecast for the largest cloud objects, which indicates that these objects have centers of mass that are closer to their matched forecast objects in the FH1 forecast. Because this improvement in the centroid distance interest score occurs for the largest objects, it has a greater impact on the MSS because the MSS is an area-weighted calculation.

Differences in the percent of occurrences in the boundary distance interest scores are shown in Fig. 8. Overall, the highest interest values are associated with FH0, which indicates that the model initialization is more accurate than the FH1 forecast for this particular attribute. As was seen in the centroid distance scores in Fig. 7, lower interest values are more common for the small cloud objects in the FH1 forecasts. This is due to greater displacement between the

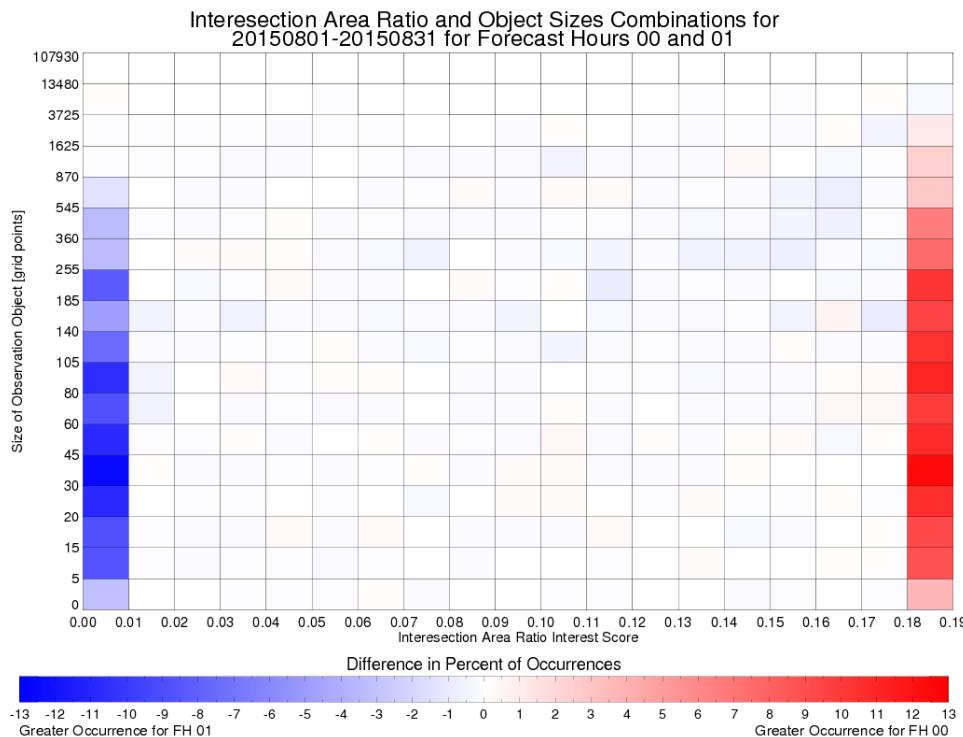
small cloud objects. However, unlike the centroid distance interest scores, there are only very small differences for the larger cloud objects. Because the boundary distance is the minimum distance between the edges of the observed and forecast cloud objects, larger objects need larger displacement between cloud objects for no overlapping of objects to occur. The difference in percent of occurrences for the intersection area ratio interest score (Fig. 9) is very similar to the boundary distance interest scores for this same reason. As the displacement between objects increases, the size of the overlapping area, and thus the intersection area ratio interest score, decreases. However, matched pairs with larger observed objects are still overlapping, and no changes are evident between FH0 and FH1 for the larger objects.



*Figure 7. Percent difference in the occurrence of MODE centroid distance interest scores between FH0 and FH1 during August 2015 plotted as a function of MODE interest score along the x-axis and observed object size along the y-axis. Blue (red) colors indicate that the given centroid distance interest score and observed object size combination occurs more (less) frequently in the FH1 forecast than in the FH0 initialization.*



*Figure 8.* Same as Fig. 7 except for the MODE boundary distance attribute.



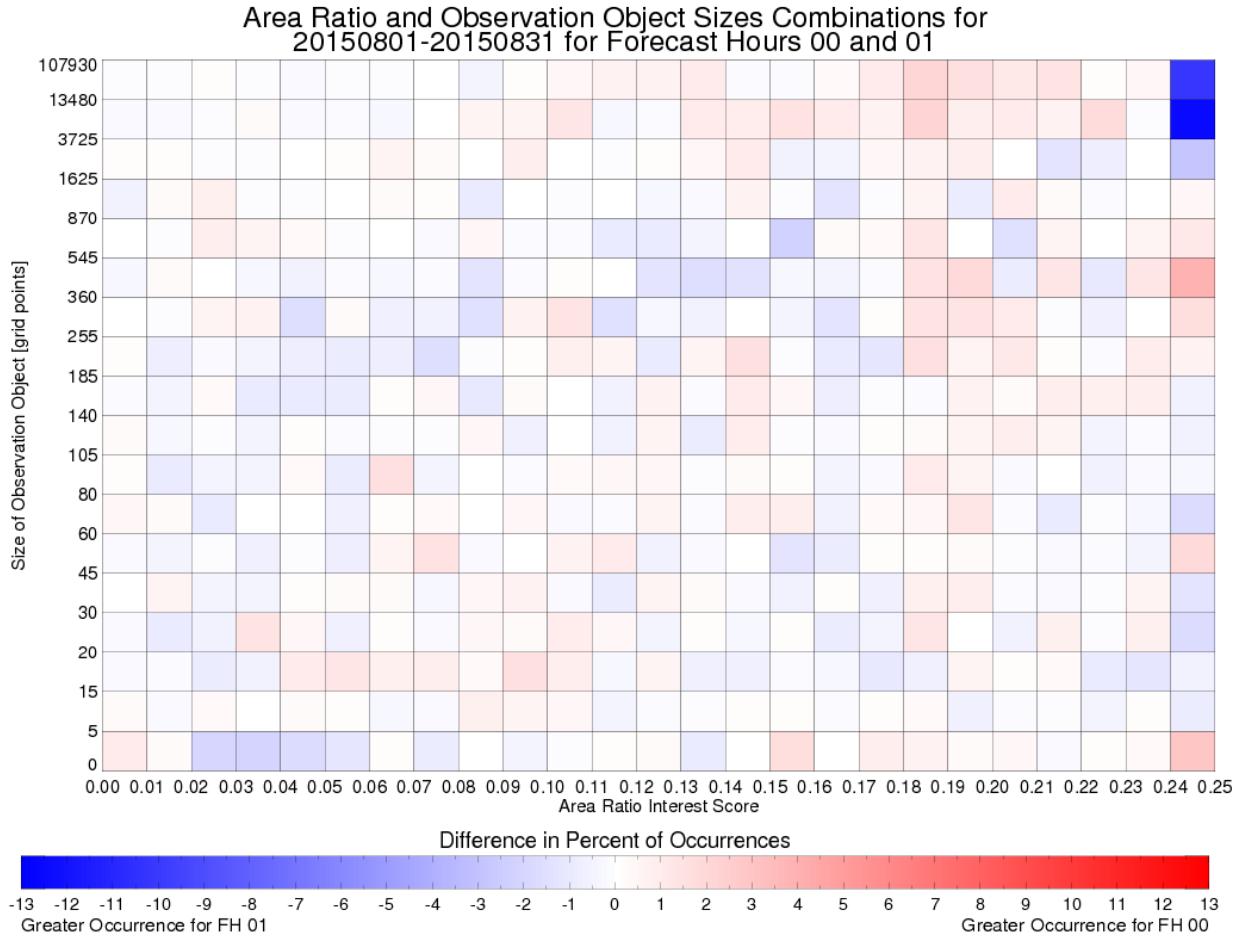
*Figure 9.* Same as Fig. 7 except for the MODE intersection area ratio attribute.

Finally, the difference in the percent of occurrence for the area ratio attribute is shown in Fig. 10. Overall, the differences between FH0 and FH1 are relatively small for this attribute. An important exception, however, is the largest observed objects, where higher area ratio interest scores are more associated with the FH1 forecast than the FH0 initialization. Because the MSS is an area-weighted calculation, the much better representation of larger cloud objects in the FH1 forecast will have a larger impact than the smaller objects. This increase in the area ratio interest score for FH1 indicates that the size of the cloud objects in the 1-h forecast better represents the observed cloud objects than the 0-h initialization. Because the observed object sizes remain unchanged, it is assumed this better representation of large observed objects is a result of their matched forecast object increasing in size. This would explain how the intersection area ratio remains unchanged between FH0 and FH1 for larger cloud objects. In addition, the center of mass between the objects could decrease even though displacement occurs because forecast objects are larger.

## 6. Summary

The primary goal of this project was to learn how to use MODE to assess the accuracy of HRRRx forecasts through comparisons of observed and forecast cloud objects occurring in the upper troposphere. Emphasis was placed not only on using MODE output, but also manipulating the MODE output to provide new insights into forecast accuracy. One outcome of our work with the DTC is the creation of the MSS. The MSS uses the MODE interest values to create a single value representing forecast accuracy, which promotes a quicker comparison between multiple forecasts valid at the same time. In addition, MODE output is used to regenerate the attribute

interest scores (which are not readily available in the current MODE output) to investigate why the 1-h HRRRx forecast is more accurate than the 0-h initialization. It is found that, even though displacement errors between the forecast and observed cloud objects increase between FH0 and FH1, that the sizes of the 1-h forecast objects better represent the observed cloud objects. This improvement for the largest cloud objects has a more positive impact on the MSS and forecast accuracy than did the degradations that occurred for the smaller cloud objects. Finally, these results were also shared with HRRR model developers at ESRL and the EMC at the end of the project. These trips provided valuable opportunities to interact with the model developers and to contribute to their model assessment and development activities.



*Figure 10. Same as Fig. 7 except for the MODE area ratio attribute.*

## 7. Publications resulting from this project

Griffin, S. M., J. A. Otkin, C. M. Rozoff, J. M. Sieglaff, L. M. Cronce, and C. R. Alexander, 2016: Methods for comparing simulated and observed satellite infrared brightness temperatures and what do they tell us? Submitted to *Wea. Forecasting*.

Griffin, S. M., J. A. Otkin, C. M. Rozoff, J. M. Sieglaff, L. M. Cronce, C. R. Alexander, R. Bullock, J. Halley-Gotway, T. Jensen, and J. Wolff, 2016: Seasonal analysis of cloud objects in the High Resolution Rapid Refresh (HRRR) model using object-based verification. In preparation for submission to *Mon. Wea. Rev.*

## 8. References

Benjamin, S., S. Weygandt, J. Brown, M. Hu, C. Alexander, T. Smirnova, J. Olson, E. James, D. Dowell, G. Grell, H. Lin, S. Peckham, T. Smith, W. Moninger, J. Kenyon, and G. Manikin, 2016: A North American Hourly Assimilation and Model Forecast Cycle: The Rapid Refresh. *Mon. Wea. Rev.* doi:10.1175/MWR-D-15-0242.1.

Cai H. and R. E. Dumais Jr., 2015: Object-Based Evaluation of a Numerical Weather Prediction Model's Performance through Forecast Storm Characteristic Analysis. *Wea. Forecasting*, **30**, 1451–1468.

Cintineo, R., J. A. Otkin, F. Kong, and M. Xue, 2014: Evaluating the accuracy of planetary boundary layer and cloud microphysical parameterization schemes in a convection-permitting ensemble using synthetic GOES-13 satellite observations. *Mon. Wea. Rev.*, **142**, 107-124.

Davis, C. A., B. G. Brown, and R. G. Bullock, 2006a: Object-based verification of precipitation forecasts. Part I: Methodology and application to mesoscale rain areas. *Mon. Wea. Rev.*, **134**, 1772-1784.

Davis, C. A., B. G. Brown, and R. G. Bullock, 2006b: Object-based verification of precipitation forecasts. Part II: Application to convective rain systems. *Mon. Wea. Rev.*, **134**, 1785-1795.

\_\_\_\_\_, \_\_\_\_\_, \_\_\_\_\_, 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.

Developmental Testbed Center, 2014: Model Evaluation Tools Version 5.0 (METv5.0) User's Guide 5.0. [Available online at [http://www.dtcenter.org//met/users/docs/users\\_guide/MET\\_Users\\_Guide\\_v5.0.pdf](http://www.dtcenter.org//met/users/docs/users_guide/MET_Users_Guide_v5.0.pdf)]

Earth System Research Laboratory, cited 2016: Rapid Refresh. [Available online at <http://rapidrefresh.noaa.gov/>]

Griffin, S. M., J. A. Otkin, C. M. Rozoff, J. M. Sieglaff, L. M. Cronce, and C. R. Alexander, 2016: Methods for comparing simulated and observed satellite infrared brightness temperatures and what do they tell us? *Wea. Forecasting*. In Review

Han, Y., P. van Delst, Q. Liu, F. Weng, B. Yan, R. Treadon, and J. Derber, 2006: JCSDA Community Radiative Transfer Model (CRTM)—version 1. NOAA Tech. Rep. NESDIS 122, 40 pp.

Mittermaier, M. P., and R. Bullock, 2013: Using MODE to explore the spatial and temporal characteristics of cloud cover forecasts from high-resolution NWP models. *Meteorol. Appl.*, **20**, 187-196.

Otkin, J. A., D. J. Posselt, E. R. Olson, H.-L. Huang, J. E. Davies, J. Li, and C. S. Velden, 2007: Mesoscale numerical weather prediction models used in support of infrared hyperspectral measurements simulation and product algorithm development. *J. Atmospheric and Oceanic Tech.*, **24**, 585-601.

Otkin, J. A., T. J. Greenwald, J. Sieglaff, and H.-L. Huang, 2009: Validation of a large-scale simulated brightness temperature dataset using SEVIRI satellite observations. *J. Appl. Meteor. Climatol.*, **48**, 1613-1626.

Otkin, J. A., and T. J. Greenwald, 2008: Comparison of WRF model-simulated and MODIS-derived cloud data. *Mon. Wea. Rev.*, **136**, 1957-1970.

Thompson, G., M. Tewari, K. Ikeda, S. Tessendorf, C. Weeks, J. A. Otkin, and F. Kong, 2016: Explicitly-coupled cloud physics and radiation parameterizations and subsequent evaluation in WRF high-resolution convective forecasts. *Atmos. Res.*, **168**, 92-104.