Using Object-Based Verification Methods and Satellite Infrared Brightness Temperatures to Assess the Accuracy of High-Resolution Models

Sarah Griffin, ***Jason Otkin***, *Justin Sieglaff, Lee Cronce, and Sharon Nebuda*
Cooperative Institute for Meteorological Satellite Studies, University of Wisconsin-Madison
Tara L. Jensen, and Eric Gilleland
National Center for Atmospheric Research/Research Applications Laboratory and Developmental Testbed Center
Patrick S. Skinner
Cooperative Institute for Mesoscale Meteorological Studies
Timothy Supinie and Ming Xue
Center for Analysis and Prediction of Storms

This research is supported by the NOAA Joint Technology Transfer Initiative and NOAA GOES-R Risk Reduction programs.

Utility of Satellite Brightness Temperatures

- Sensors provide detailed information about the state of the atmosphere and land/ocean surfaces
    - Obtained from visible, shortwave, infrared, and microwave bands
- Our studies have focused on all-sky infrared brightness temperatures because they provide valuable information about clouds and water vapor, both of which are susceptible to large errors in NWP model forecasts
- Geostationary satellites provide routine coverage over large areas, whereas polar-orbiting satellites provide global coverage, but with less frequent updates

Simulated Satellite Brightness Temperatures

- Must convert the model forecasts into simulated, or model-equivalent, brightness temperatures using a forward radiative transfer model (RTM)

- NWP model fields used by the RTM typically include T, q_v, T_{skin}, 10-m wind speed, and the mixing ratios and effective diameters for five hydrometeor species (cloud water, rain water, ice, snow, and graupel)

- Preferable to leave all 3-d model fields on the model's vertical grid

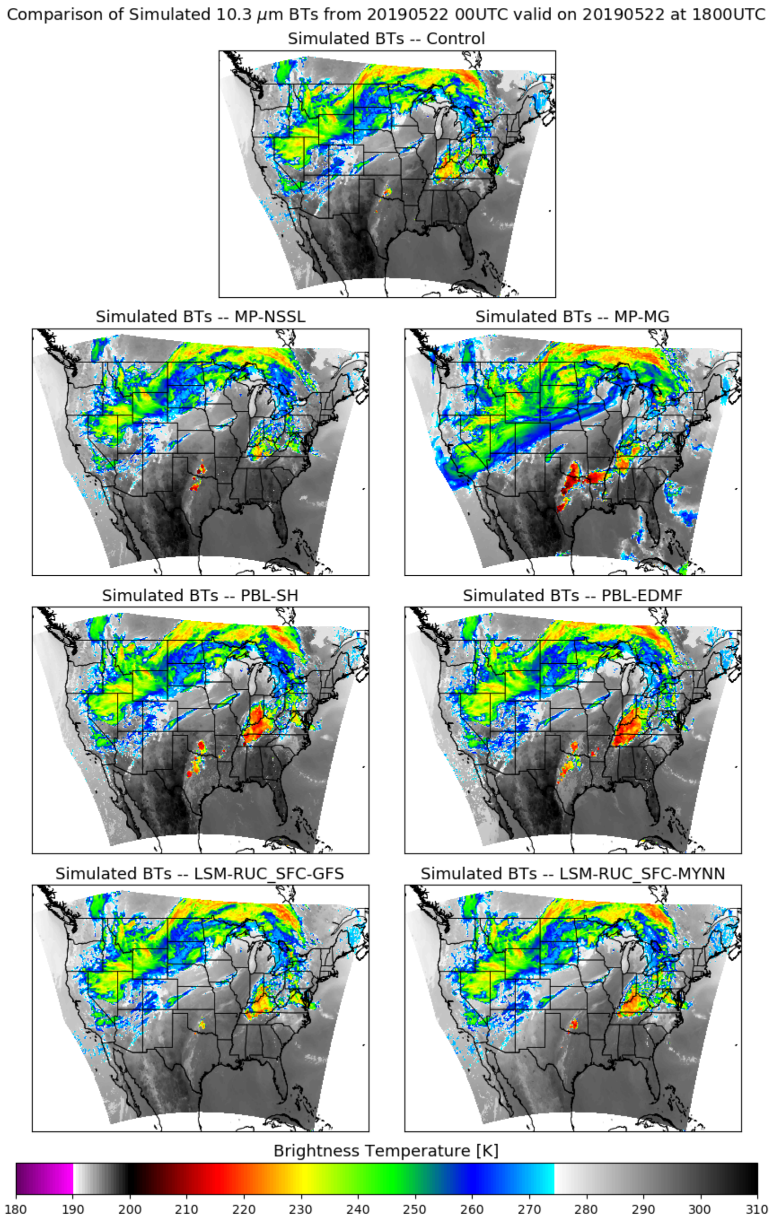
- Important to ensure that the effective particle diameters are computed correctly for each hydrometeor species

    - Should be computed based on assumptions made by a given microphysical scheme (particle density, particle size distribution, etc.)

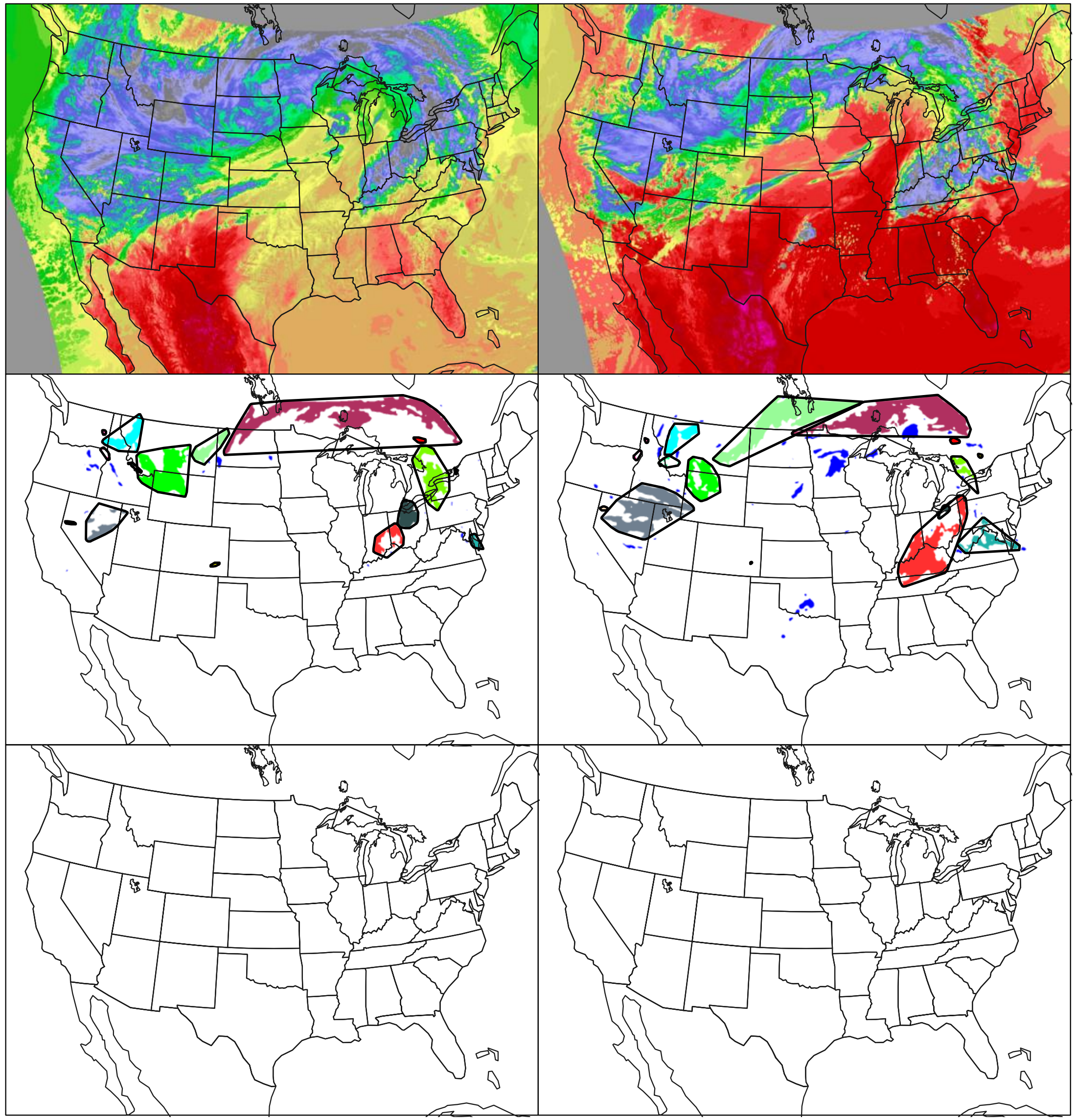
Model Configurations: FV3-LAM

Name	Microphysics Scheme	PBL Scheme	Surface Layer	LSM
Control	Thompson	MYNN	GFS	Noah
MP-NSSL	NSSL	MYNN	GFS	Noah
MP-MG	Morrison-Gettelman	MYNN	GFS	Noah
PBL-SH	Thompson	Shin-Hong	GFS	Noah
PBL-EDMF	Thompson	EDMF	GFS	Noah
LSM-RUC_SFC-GFS	Thompson	MYNN	GFS	RUC
LSM-RUC_SFC-MYNN	Thompson	MYNN	MYNN	RUC

- Simulations run during 2019 HWT Spring Experiment at OU-CAPS; 20 forecasts in total; 60 hours long, initiated at 00 UTC



Method for Object-Based Diagnostic Evaluation (MODE)



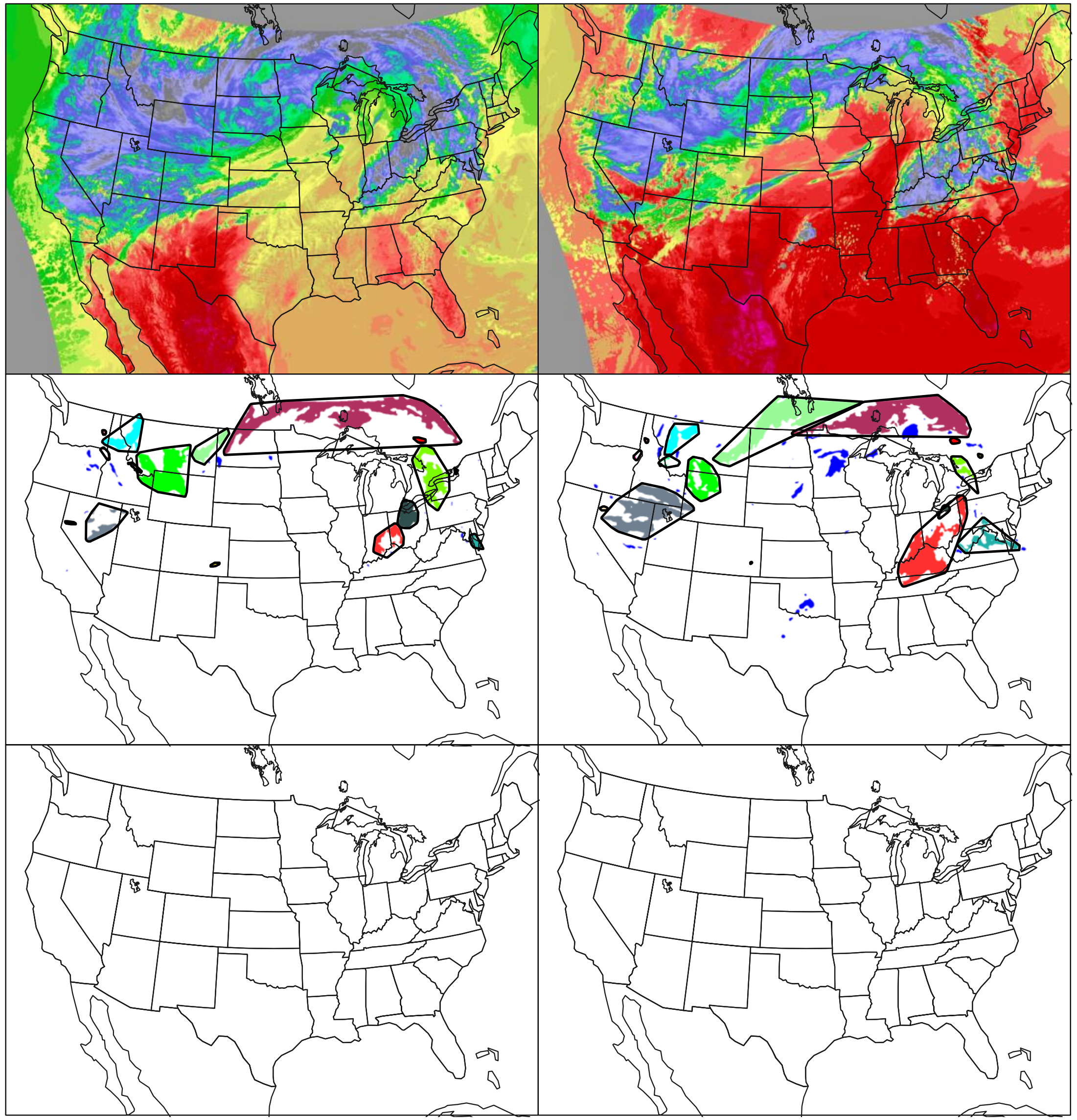
1. Identify objects in forecast and observed fields.
2. Identify various object attributes for each object, such as location and size.
3. Match the forecast and observed cloud objects.
4. Output attributes for individual objects, such as location and size, and matched object pairs, such as the distance between object centers, ratio of object sizes, and overall interest score describing the "goodness" of the match for assessment.

Method for Object-Based Diagnostic Evaluation (MODE)

Interest Scores: similarity between matching forecast and observation MODE objects

Object Pair Attribute	User-Defined Weight (%)	Description
centroid_dist	4 (25.0)	Distance between objects' "center 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 observation objects' areas (or its reciprocal, whichever yields a lower value)
int_area_ratio	3 (18.75)	Ratio of the objects' intersection area to the lesser of the observation or forecast area (whichever yields a lower value)

Method for Object-Based Diagnostic Evaluation (MODE)



Clusters: one or more observed objects matched with one or more forecast objects

- Must have an interest score > 0.65 to be used in our study
- Useful when analyzing matched object pairs, as otherwise smaller objects might not have a match and skew statistics
- Examples:
  - Gray objects over Nevada
  - Green objects over Ontario, Canada

Methodology

1. Object-based analysis

Object-based Threat Score (OTS) :

$$OTS=\frac{1}{A_f+A_o}\left[\sum_{p=1}^{P} I^p(a_f^p+a_o^p)\right]$$

A_f and A_o : Area of all forecasted and observed objects.

P : number of matched simulated and observation object pairs

I^p : interest score between the matched simulated and observation object

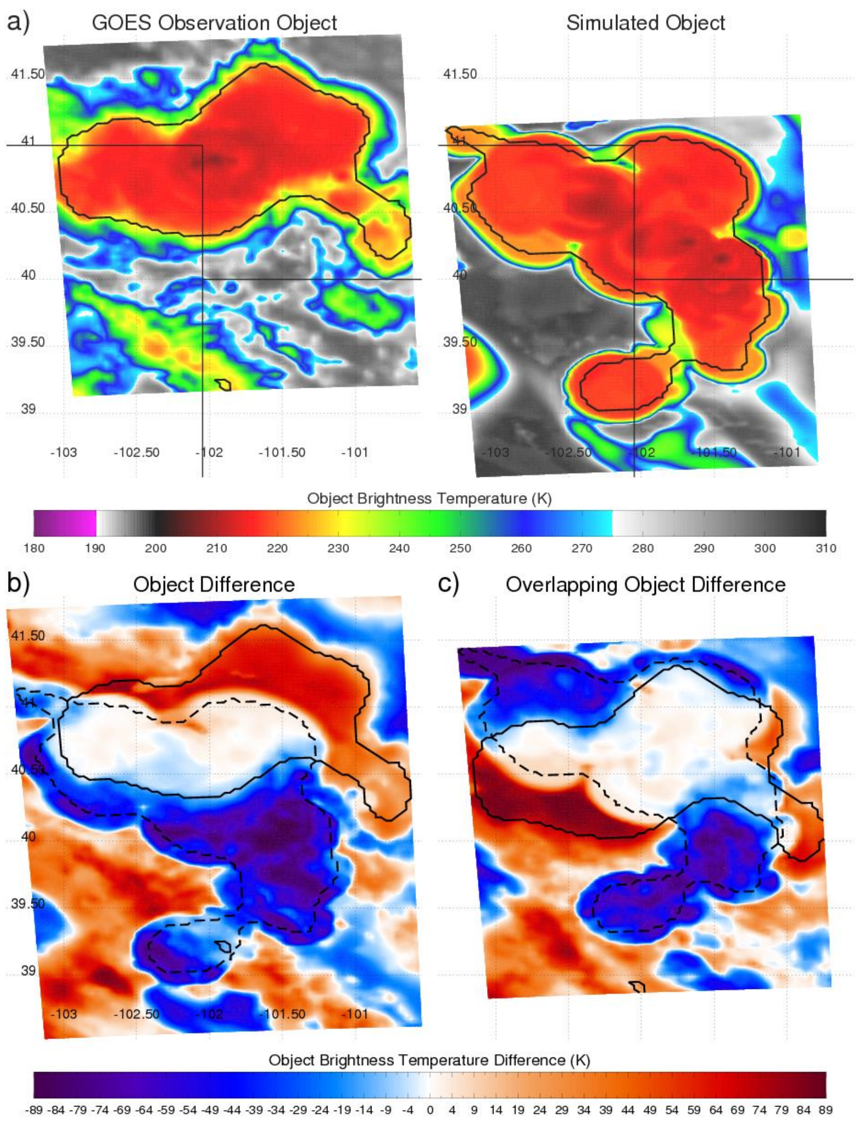
a_f^p and a_o^p : areas of the forecast and observation objects in the matched pair

2. Pixel-based analysis

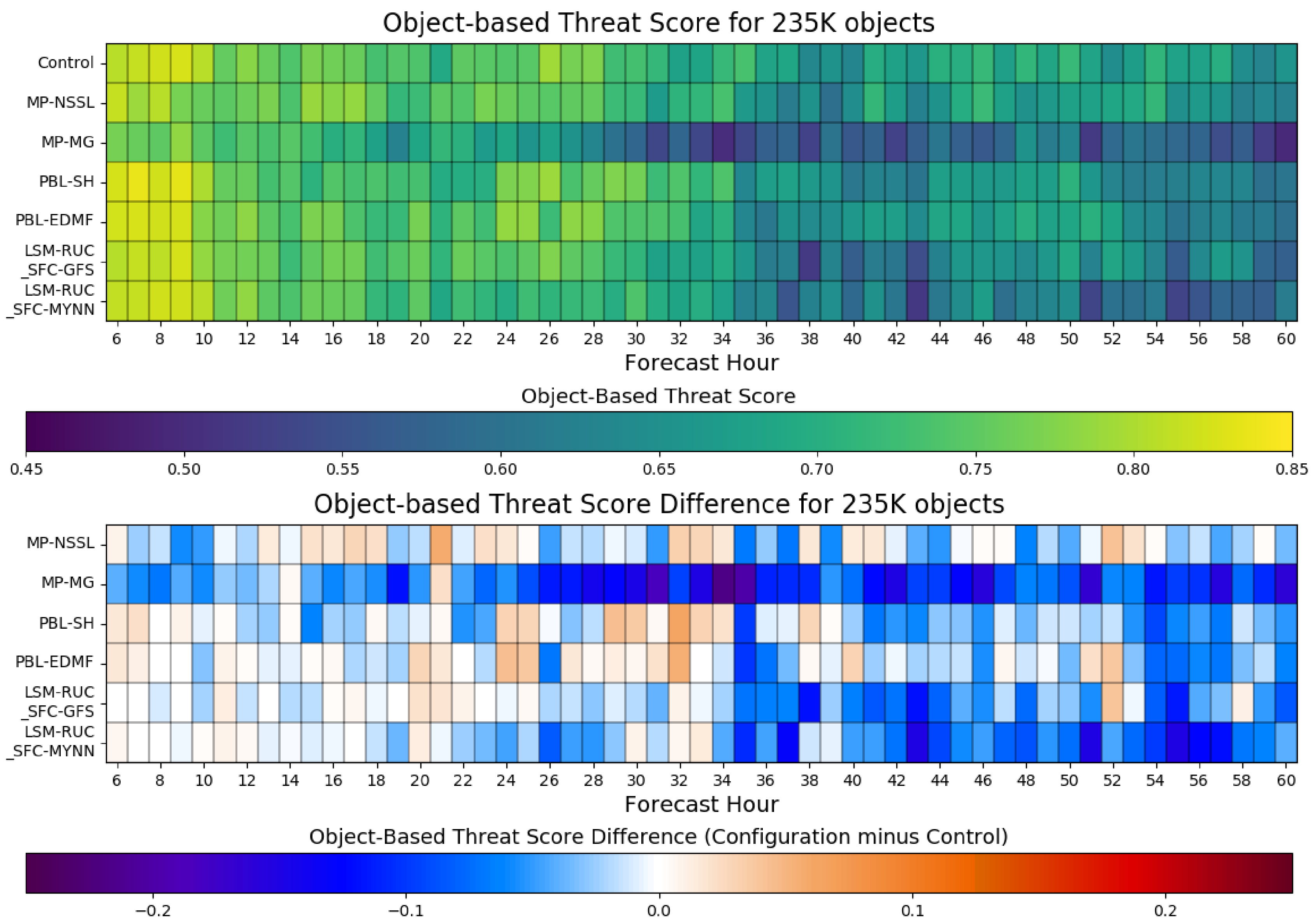
Mean Absolute Error (MAE): $MAE=\frac{1}{N}\sum_{i=1}^{N}|F_i - O_i|$

Mean Bias Error (MBE): $MBE = \frac{1}{N}\sum_{i=1}^{N}(F_i - O_i)$

F and O : forecast and observation BTs

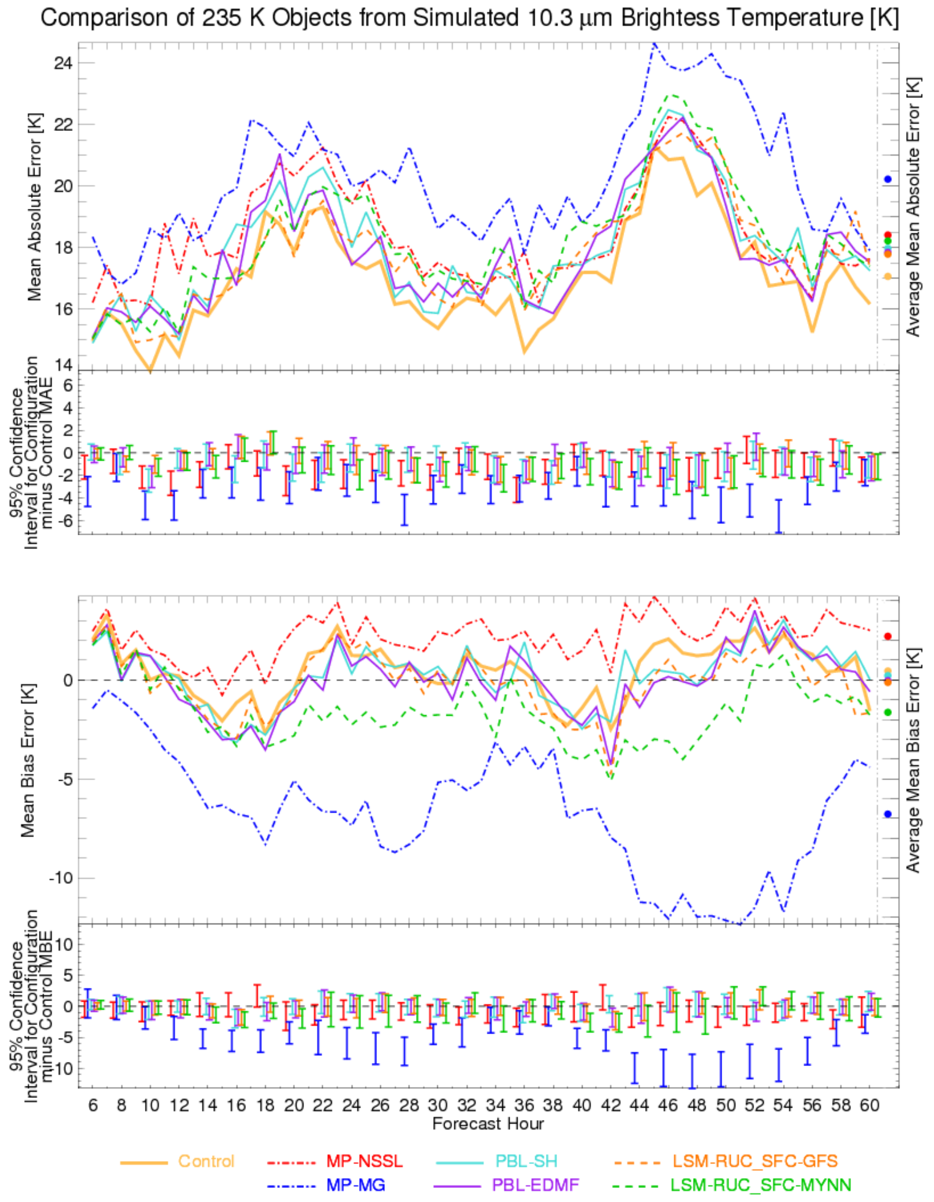


Object-Based Threat Score



- Control has the highest average OTS.
- MP-MG has the lowest average OTS.
- LSM-RUC_SFC-MYNN has the steepest decline in OTS by forecast hour.
  - Due to increased number of objects
- Parameterization changes have a neutral to positive impact on OTS in early FHs compared to Control.

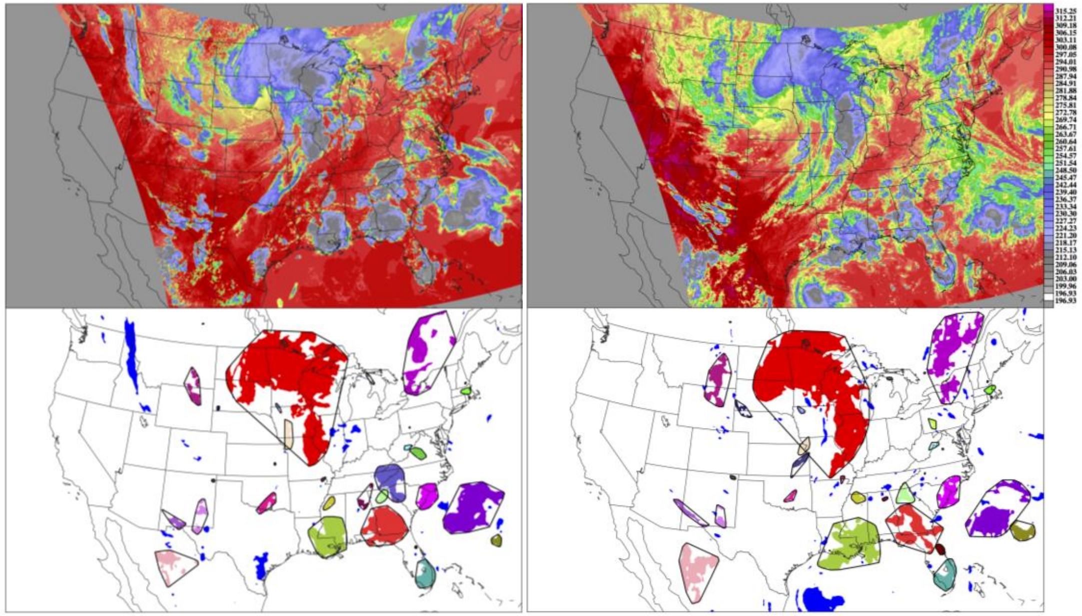
Pixel-based MAE and MBE



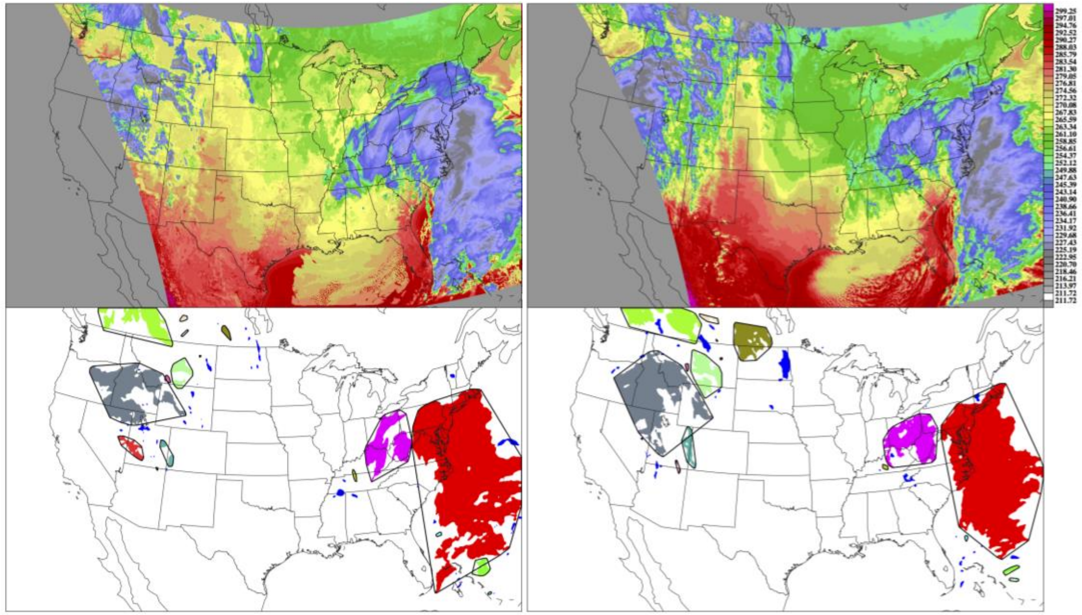
- MP-MG highest MAE, Control lowest.
  - MP-MG difference from Control statistically significant.
  - MP-NSSL next highest MAE.
- Diurnal cycle in model accuracy
  - Opposite of OTS (MAE high when OTS low)
  - Centroid distance removed for MAE.
- Changing microphysics scheme has largest impact on MBE (and MAE)
  - MP-MG low bias in object BTs
  - MP-NSSL has a high bias in object BTs.
  - MBE is correlated with an increased number of forecast object grid points compared to the observation object.

Satellite-Based Verification of the HRRR Model

4-h HRRR from 1800 UTC on 20150818 GOES from 2200 UTC on 20150818



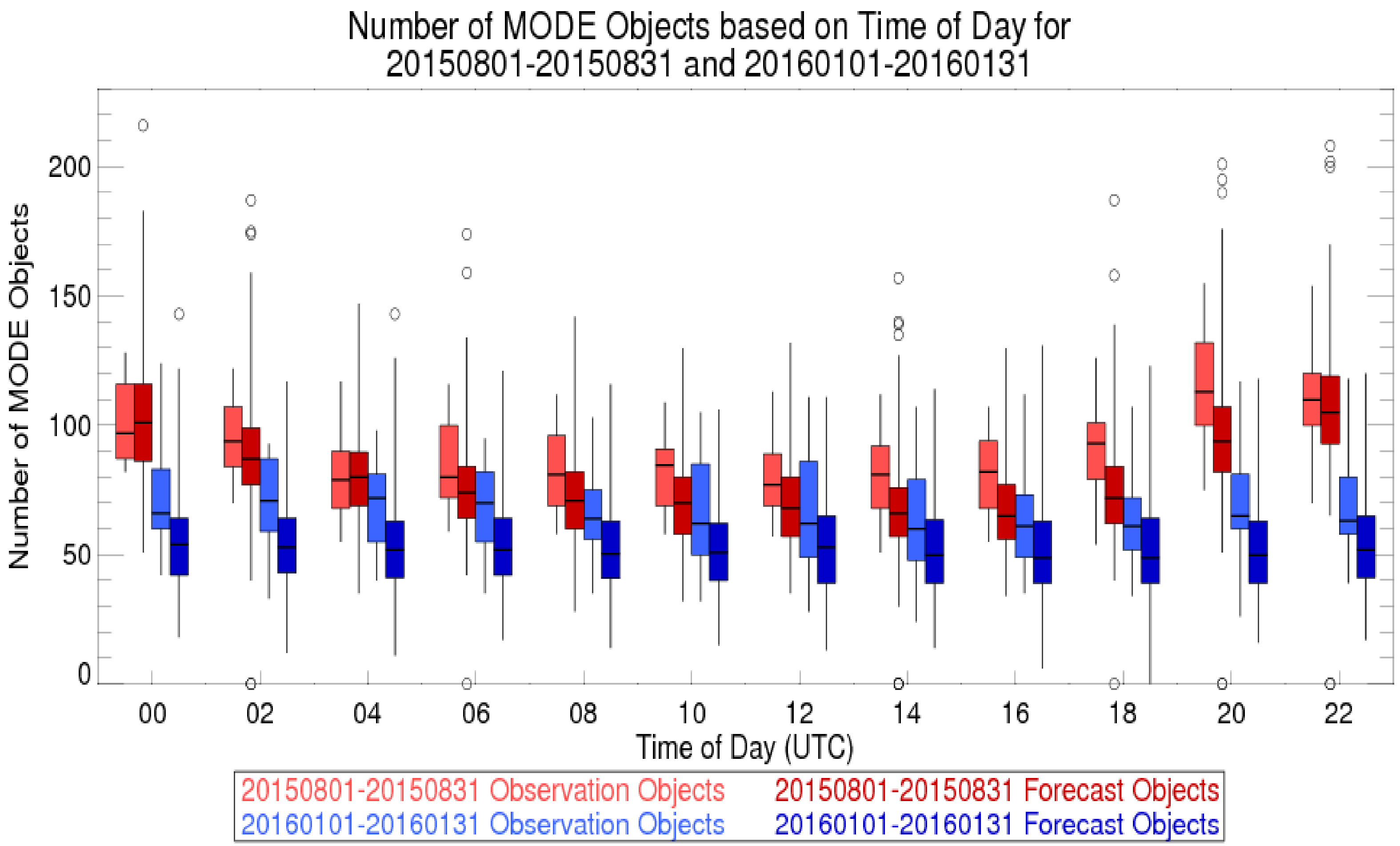
2-h HRRR from 2100 UTC on 20160122 GOES from 2300 UTC on 20160122



- **Summer Example**: Forecast hours 0-24 from HRRR initializations from August 1-31, 2015

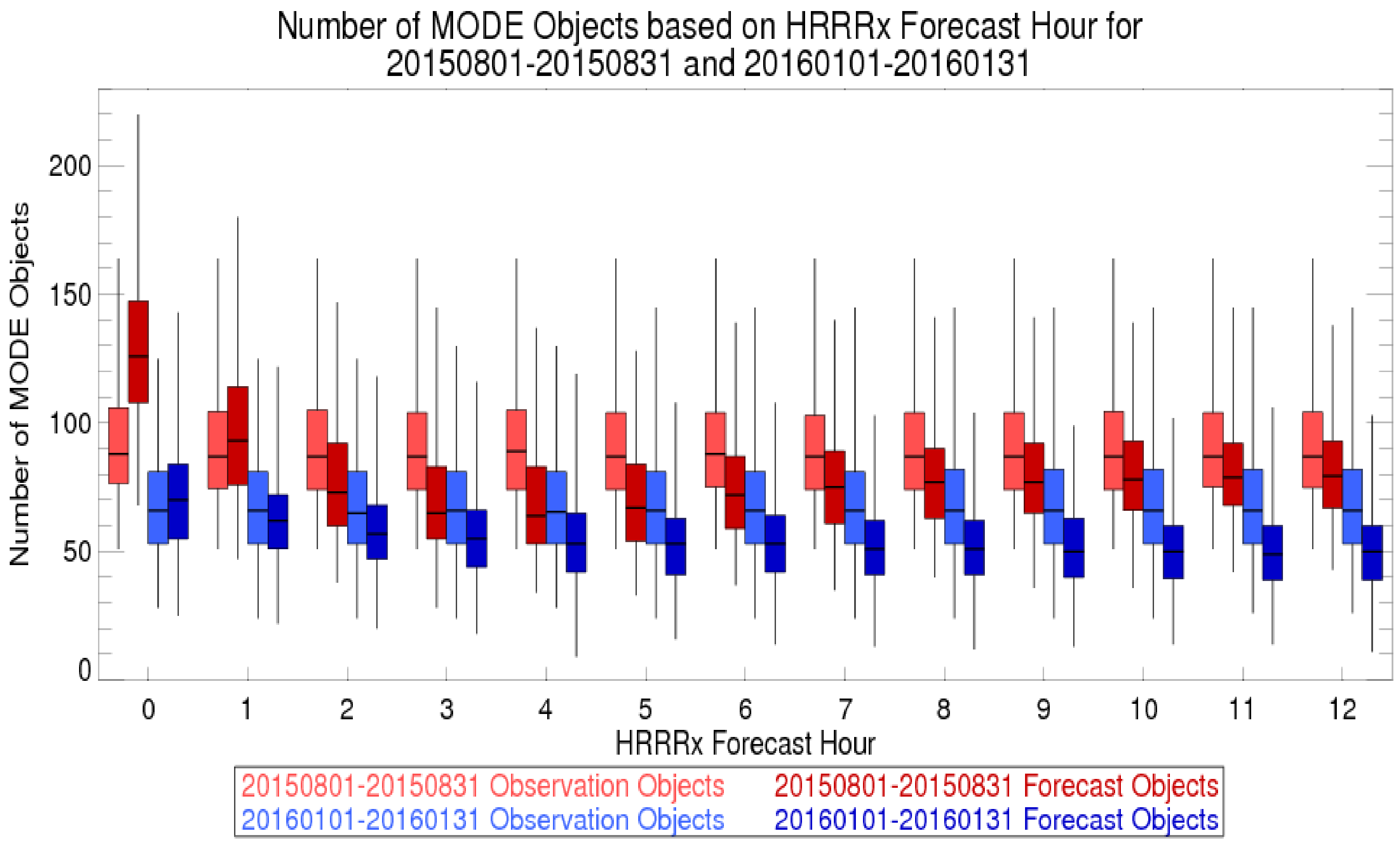
- **Winter Example**: Forecast hours 0-24 from HRRR initializations from January 1-31, 2016

Number of MODE Objects – Function of Time of Day



- MODE identifies more cloud objects during August
  - Average cloud object is smaller (larger) during August (January)
- Diurnal cycle is much larger during August
  - Minimum (maximum) near 12 UTC (20 UTC)

Number of MODE Objects – Function of Forecast Hour



- Too many forecast cloud objects during August for forecast hours 0-1
    - Indicates cloud objects are too small in HRRR initializations
- More observed cloud objects than HRRR forecast objects overall
- Steady drift toward fewer forecast cloud objects during January

Thank you for your attention!

Griffin, S. M., J. A. Otkin, G. Thompson, M. Frediani, and J. Berner, 2019: Assessing the impact of stochastic cloud microphysics using GOES-16 infrared brightness temperatures. In preparation for submission to *Mon. Wea. Rev.*

Griffin, S. M., J. A. Otkin, C. M. Rozoff, J. M. Sieglaff, L. M. Cronce, and C. R. Alexander, 2017: Seasonal analysis of cloud objects in the High Resolution Rapid Refresh (HRRR) model using object-based verification. *Mon. Wea. Rev.*, **56**, 2317-2334.

Griffin, S. M., J. A. Otkin, C. M. Rozoff, J. M. Sieglaff, L. M. Cronce, and C. R. Alexander, 2017: Methods for comparing simulated and observed satellite infrared brightness temperatures and what do they tell us? *Wea. Forecasting*, **32**, 5-25.

Otkin, J. A., W. E. Lewis, A. Lenzen, B. McNoldy, and S. Majumdar, 2017: Assessing the accuracy of the cloud and water vapor fields in the Hurricane WRF (HWRF) model using satellite infrared brightness temperatures. *Mon. Wea. Rev.*, **145**, 2027-2046.