Using Machine Learning to Produce Watch-to-Warning Severe Weather Guidance from the Warn-on-Forecast System

Sam Varga[1,2,3], Montgomery L. Flora[2,3], Corey K. Potvin[3,1]

[1]School of Meteorology, University of Oklahoma, [2]Cooperative Institute for Severe and High-Impact Weather Research and Operations, [3] NOAA/OAR/National Severe Storms Laboratory

OVERVIEW

- Severe weather forecasts are impeded by limited NWP guidance at Watch-to-Warning lead times (0-6 hours)
- Prior studies (Clark & Loken, 2022; Flora et al., 2021) show machine learning (ML) can produce skillful guidance for the NSSL Warn-on-Forecast System (WoFS) at lead times of 0-3 hours
- We trained ML models that predict probability of severe wind, hail, & tornadoes from WoFS output at lead times of 2-6 hours
- ML yields substantial improvements over rigorous baselines

GOAL

Evaluate the performance of ML algorithms and feature subsets for the prediction of severe convective weather hazards from the WoFS at lead times of 2-6 hours

METHODS

WoFS output from the 2018-2021 Spring Forecasting Experiments (644 initializations) are used. WoFS fields are coarsened to 9 km then time-composited over the 2-6 hour window. Three sizes of spatial filter are applied to create three sets of predictors, after which ensemble statistics are calculated. The baselines (BL) are neighborhood maximum ensemble probabilities of 2-5 km updraft helicity, 80-meter wind speed, and HAILCAST tuned for each hazard. Logistic regression (LR) and histogram-based gradient boosting trees (HGBT) are trained to predict the probability that a grid point is within 36 km of a report. Both ML & BL use date-based cross validation for hyperparameter tuning and are calibrated with isotonic regression.

RESULTS

- ML output a wider range of probabilities than BL and are generally more skillful and reliable (Fig. 1-2)
- Models using *only* environmental predictors are generally worse than the BL, while intrastorm models are better (Fig. 3)
- *ML models are more sensitive to the choice of predictor variable than the choice of predictor scale for this task (Fig. 3)*
- *ML guidance provides skillful addition to NWP products (Fig. 4)*

AMS Talk

This material is based upon work supported by the Joint Technology Transfer Initiative Program within the NOAA/OAR Weather Program Office under Award No. NA22OAR4590171.

Questions? Email svarga@ou.edu

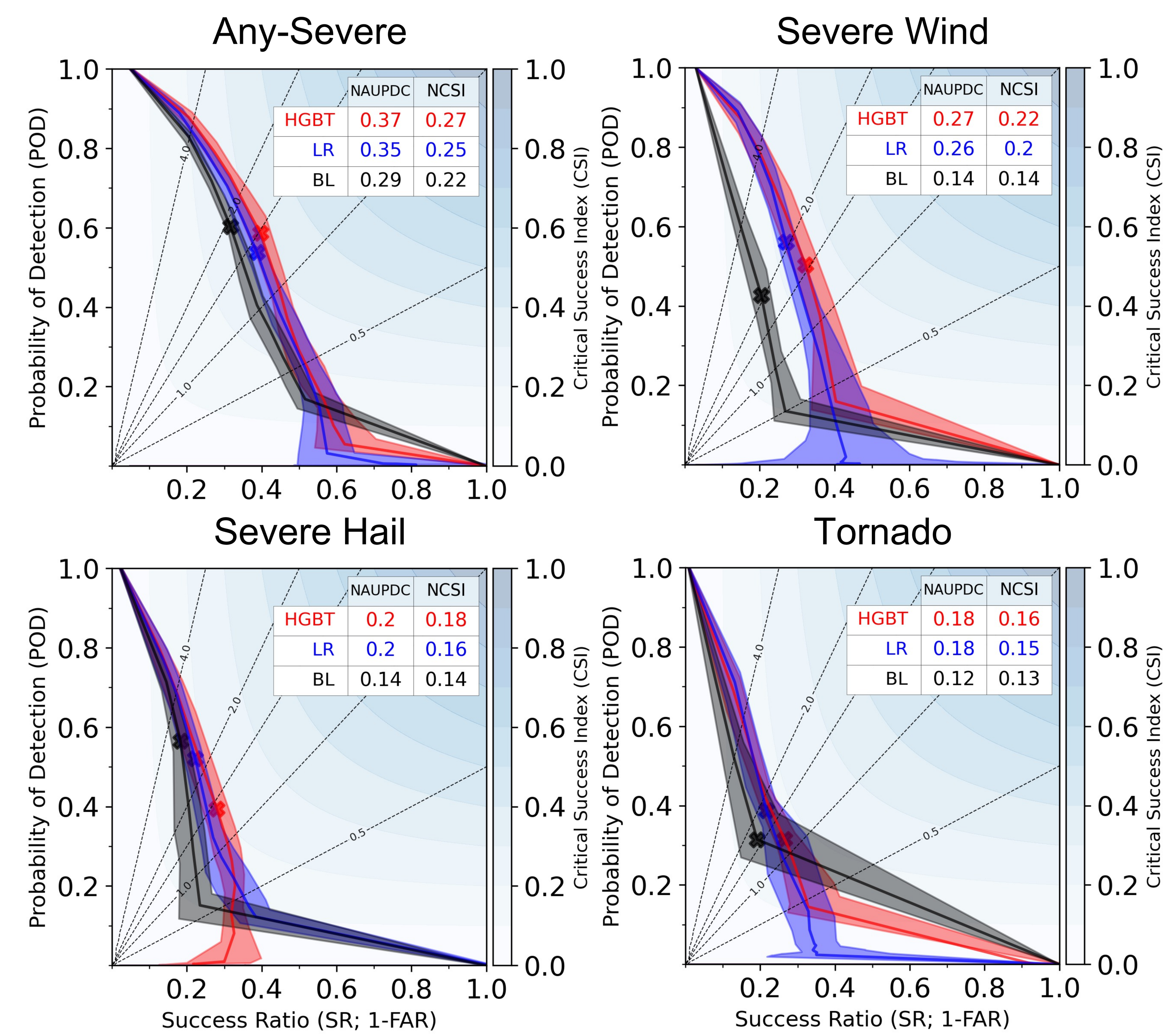

Fig. 1. Performance diagrams for ML and BL models trained to predict each hazard. While Any-Severe models have the highest skill, the severe wind ML has the largest improvement over the BL.

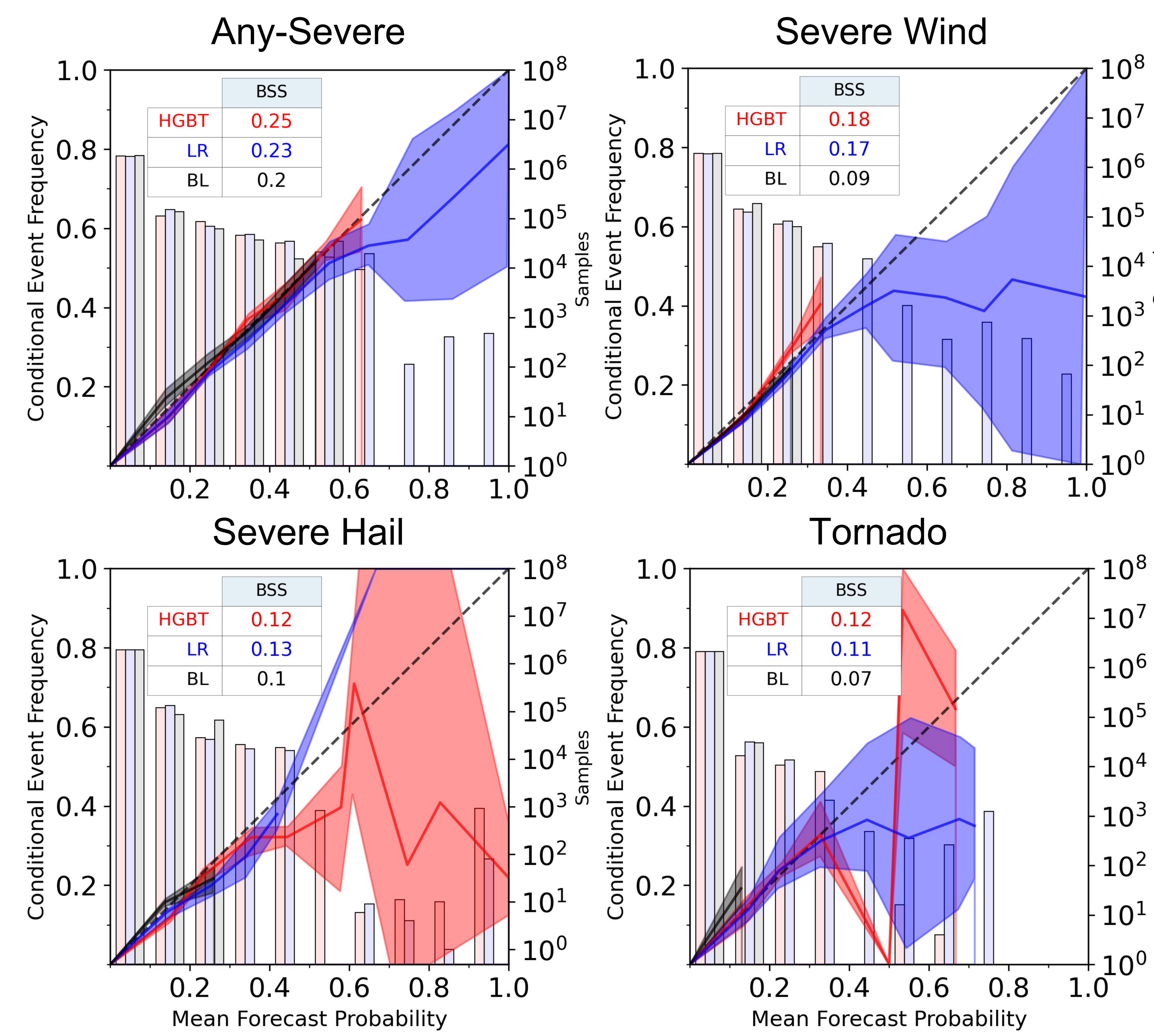

Fig. 2. Reliability diagrams and histograms of the predicted probabilities for the testing set. Shading indicates the two-sigma uncertainty interval.

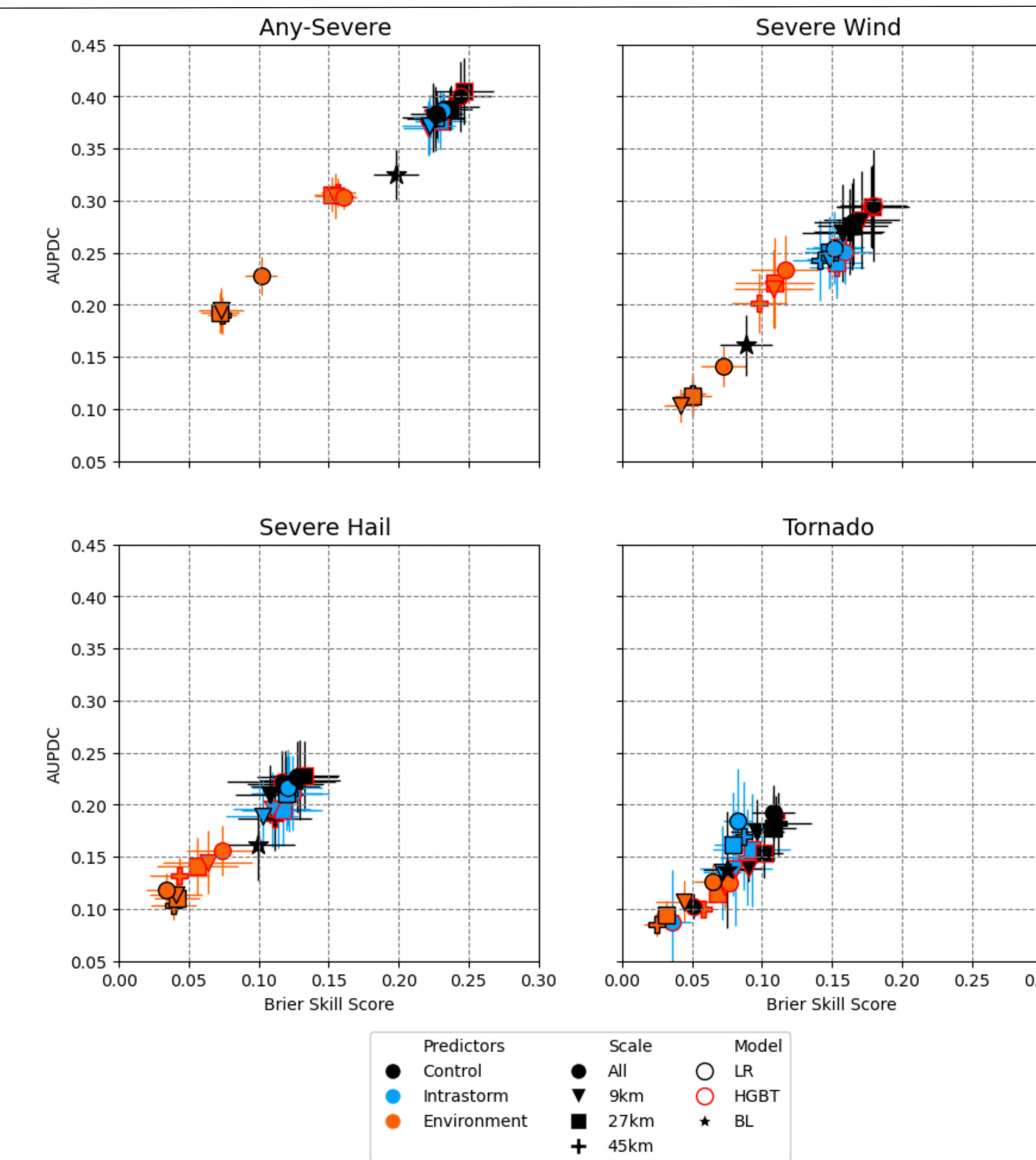

Fig. 3. Average AUPDC and BSS for ML models. The set of predictors used is specified by the color and shape of the point.

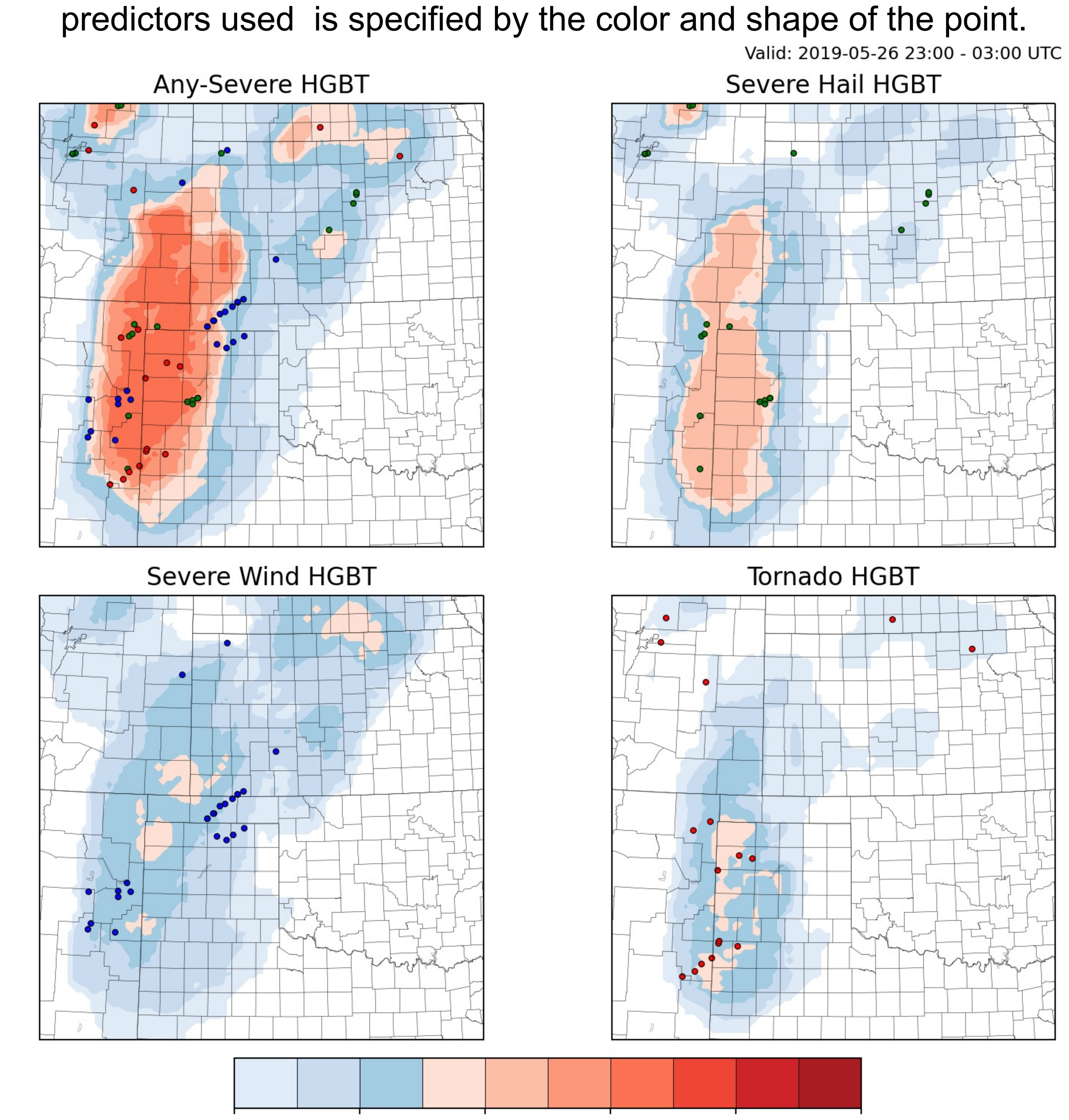

Fig. 4. Probability of hazard within 36 km in the next 2-6 hours