Exploring new techniques for WPC's Probabilistic Winter Precipitation Forecasts

Bruce Veenhuis - WPC
13 Feb. 2024

Overview

- Ensembles are unfortunately often underdispersion (over-confident) and can struggle with extremes

- There is some evidence that humans can serve as a 'safety net' in these situations
    - For example, lake effect snowfall

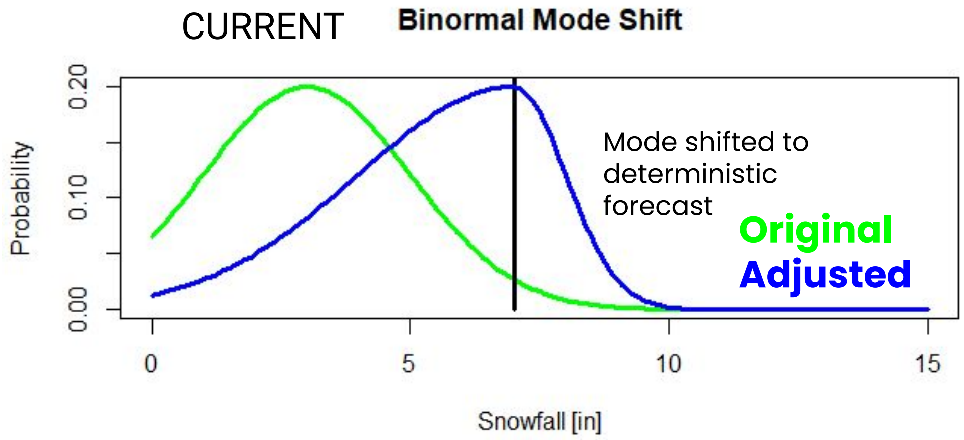
- WPC has been testing an approach that embraces BOTH objective and calibrated ensemble techniques, while also allowing the human forecast to 'nudge' the tails of the distribution.

- The revised technique shows SUBSTANTIAL improvement over existing operational WPC and NBM approaches.

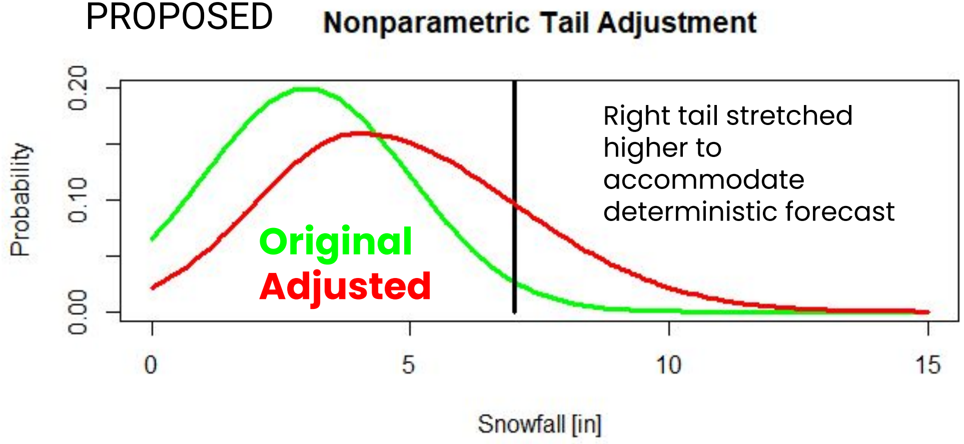
Overview - Continued

- The current operation PWPF is based on the binormal distribution
- We are testing a nonparametric method that still incorporates the deterministic forecast, but produces better reliability
- The method combines downscaled ensemble members, regression calibrated probabilities and a deterministic forecast
- This presentation will provide independent verification for the 2023-2024 winter season

CURRENT



PROPOSED

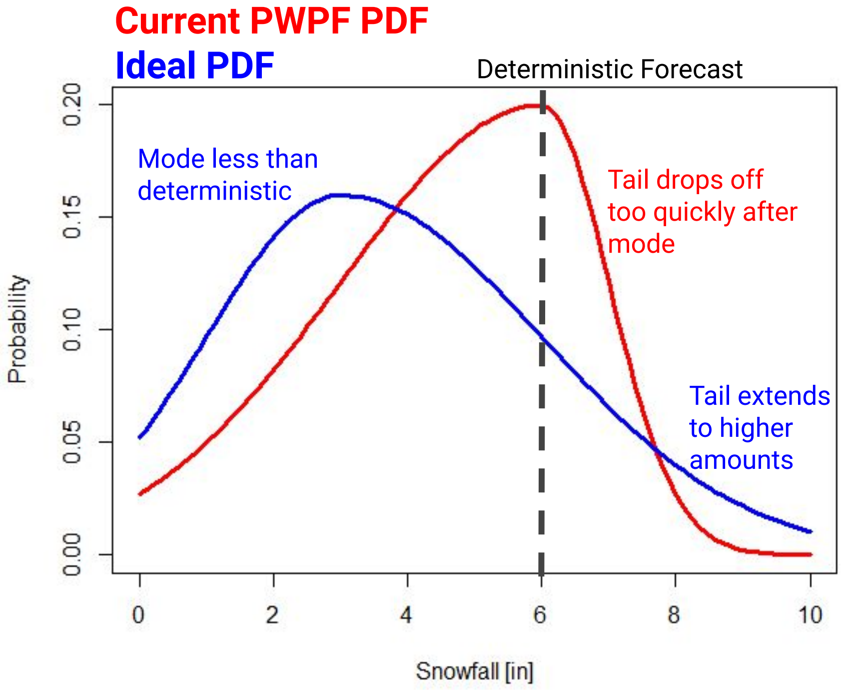


Background

- WPC's binormal methodology has generally been found to produce well calibrated 90th percentiles (at least in the bulk statistical sense), but we have often found the 10th percentile is too high leading to overconfident reliability diagrams.

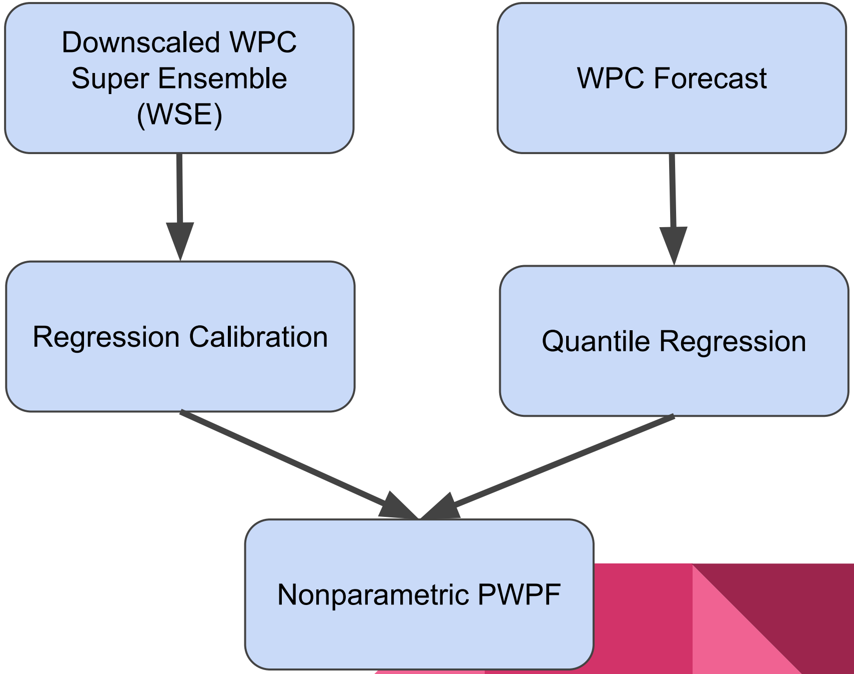
- The issue stems at least partially from the mode adjustment procedure used by the PWPF. Setting the PDF mode to equal the deterministic forecast often results in an incorrectly skewed PDF. That is, there is a heavy left tail and a sharp drop off to the right of the mode. This shape often results in little separation between the deterministic forecast and the 90th/95th percentiles.

- We have a found a distribution with right skewness and a mode somewhat less than the deterministic forecast tends to verify better

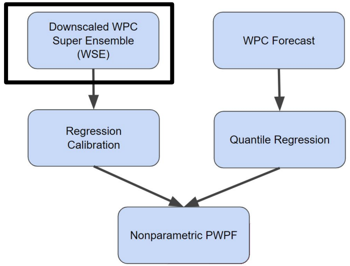


Nonparametric Method Overview

- The nonparametric approach combines probabilistic information from a few sources
- Downscaled ensemble members
- Statistically calibrated probabilities
- A deterministic forecast
- Statistical calibration is performed using a combination of multiple linear regression and quantile regression



Ensemble Member Downscaling



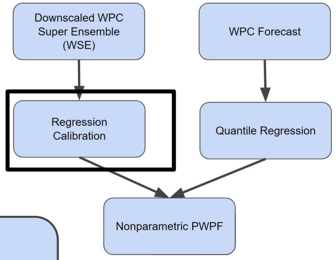
- A variety of processing is done to each ensemble member to downscale the data to 5-km. A later downscaling step produces the 2.5-km resolution PWPF in operations.

- A simple QPF PRISM downscaling is applied to the QPF of each member

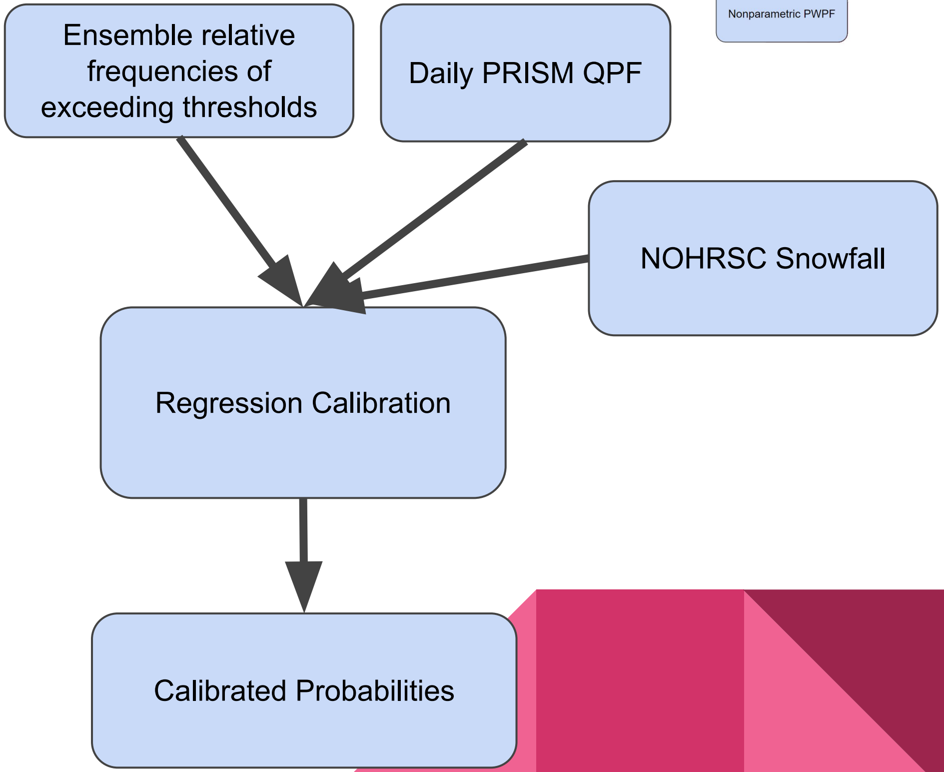
- The precipitation type grids are downscaled using a combination of snow level and downscaled 2-m temperature

- Snowfall is derived via the downscaled QPF, ptype, and an SLR (based on NBM)

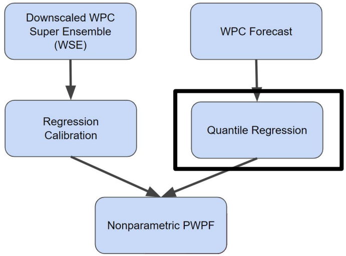
Statistically Calibrated Probabilities



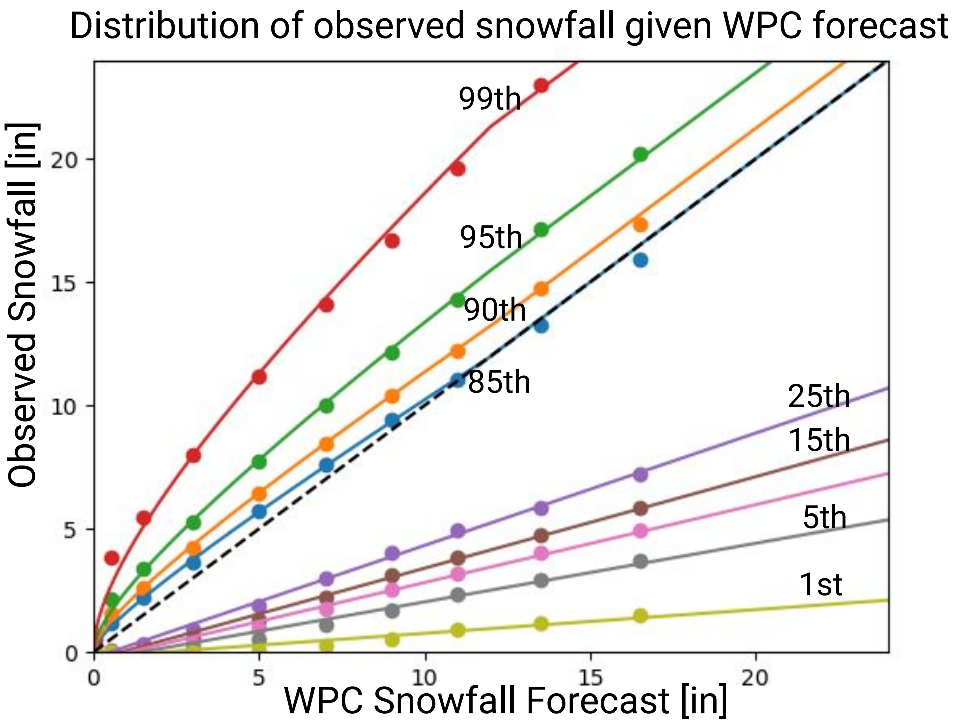
- The downscaled ensemble members are used to evaluate a set of regression equations that produced calibrated probabilities of snowfall exceeding various thresholds
- The NOHRSC snowfall exceeding various thresholds is used as a binary predictand
- The inputs to these regression equations are ensemble relative frequencies of the snowfall exceeding various thresholds and also gridpoint-specific daily PRISM QPF information
- The PRISM information is used both as an individual predictor and to form interactive predictors by multiplication with the ensemble relative frequencies
- The goal of the PRISM information is to potentially enhance snowfall over elevated terrain. This approach is different from traditional PRISM downscaling in that with the regression, PRISM only influences the result to the extent it improves the regression reduction of variance.



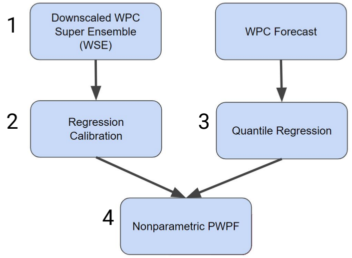
Quantile Regression with Deterministic Forecast



- A calibrated set of probabilistic information is also created using quantile regression and the deterministic forecast
- Here the input is the WPC deterministic forecast. A set of quantile regression equations are developed to provide calibrated CDF values given the deterministic forecast value.
- For simplicity, a single set of CONUS-wide equations are developed
- The goal of this step is to provide a fallback in the event the deterministic forecast is at odds with the calibrated ensemble information



Nonparametric Approach Steps



1. WPC downscales the various ensemble members of the WPC Super Ensemble (WSE)

2. The WSE and PRISM are used to evaluate the statistical regression equations

3. An initial calibrated CDF is generate from the downscaled ensemble and regression probabilities

4. The quantile regression equations are evaluated using WPC’s deterministic forecast to produce a CDF

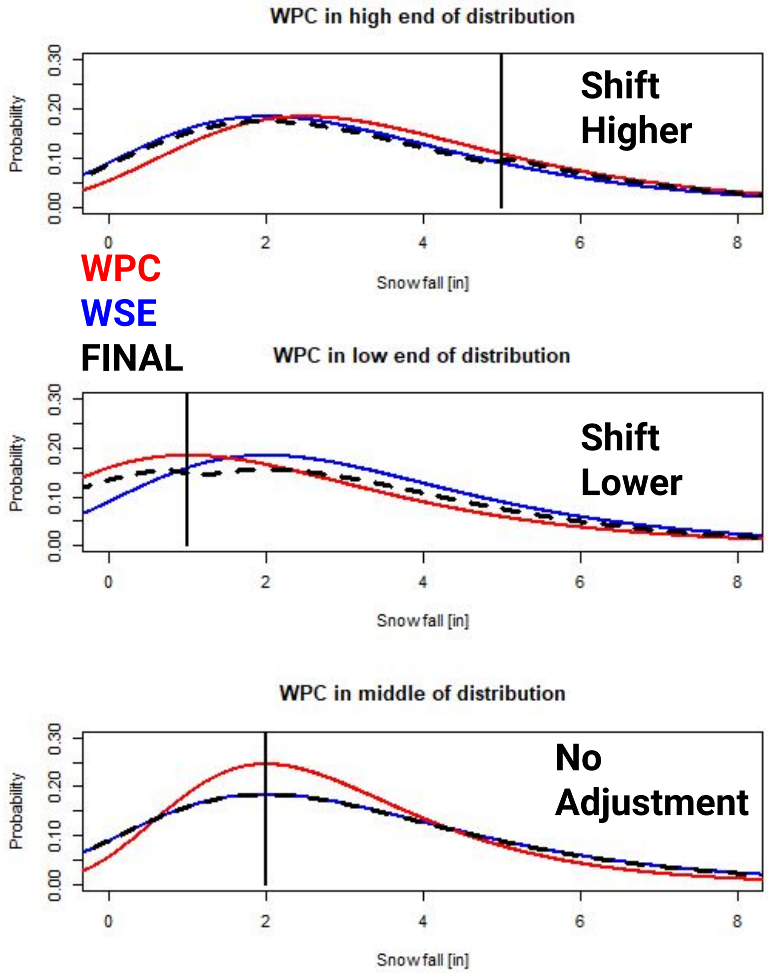
Nonparametric Approach Steps

5. When WPC is near the 85th percentile the right tail is stretched higher

6. When WPC is near the 15th percentile the left tail is stretched lower

7. If WPC is near the center of the distribution no adjustment is made

Note, for demonstration purpose these examples use parametric distributions. However, in practice the PDFs are all nonparametric and a set of percentiles is used to approximate the CDFs/PDFs.

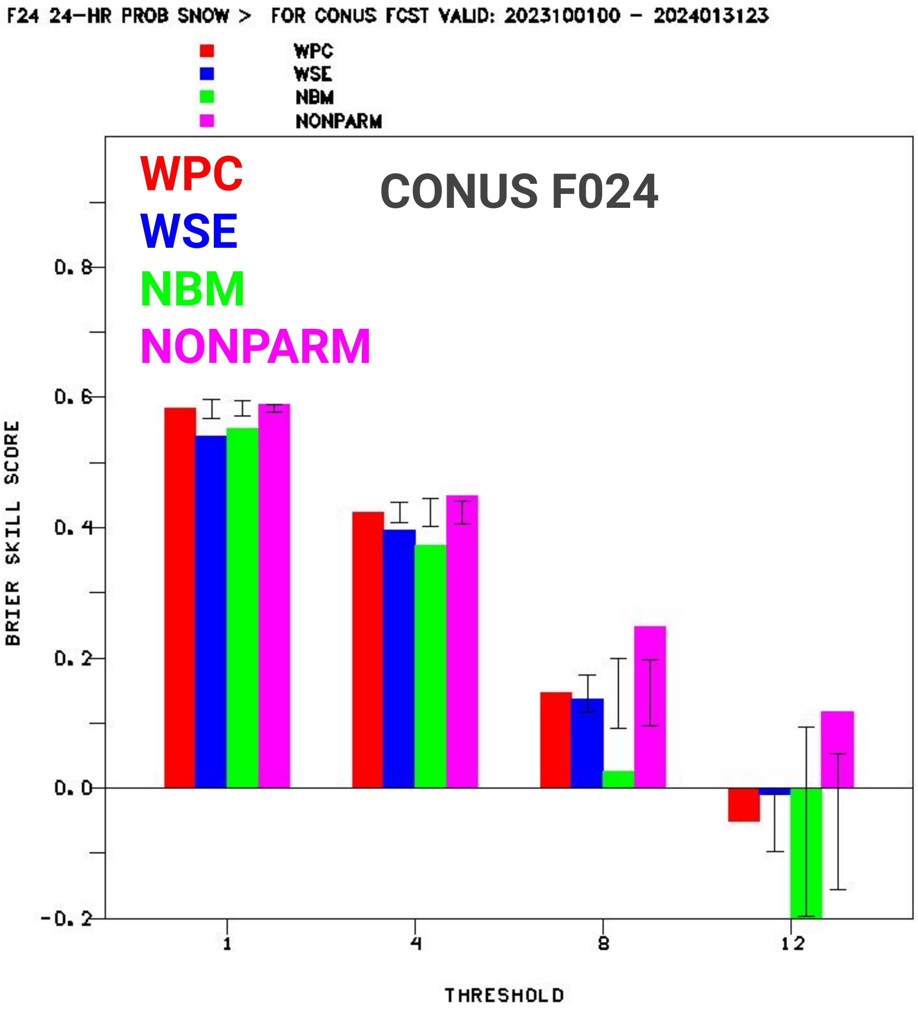


Verification Results

- All results are grid-to-grid versus the NOHRSC snowfall analysis for 24 hour durations.
- Forecasts valid 1 Oct. 2023 - 31 Jan. 2024. Two cycles per day.
- NONPARM is the new methodology
- WPC is the operational WPC PWPF
- NBM is the operational NBM PWPF

Brier Skill Score
F024 Over CONUS
1 Oct. 2023 - 31 Jan. 2024

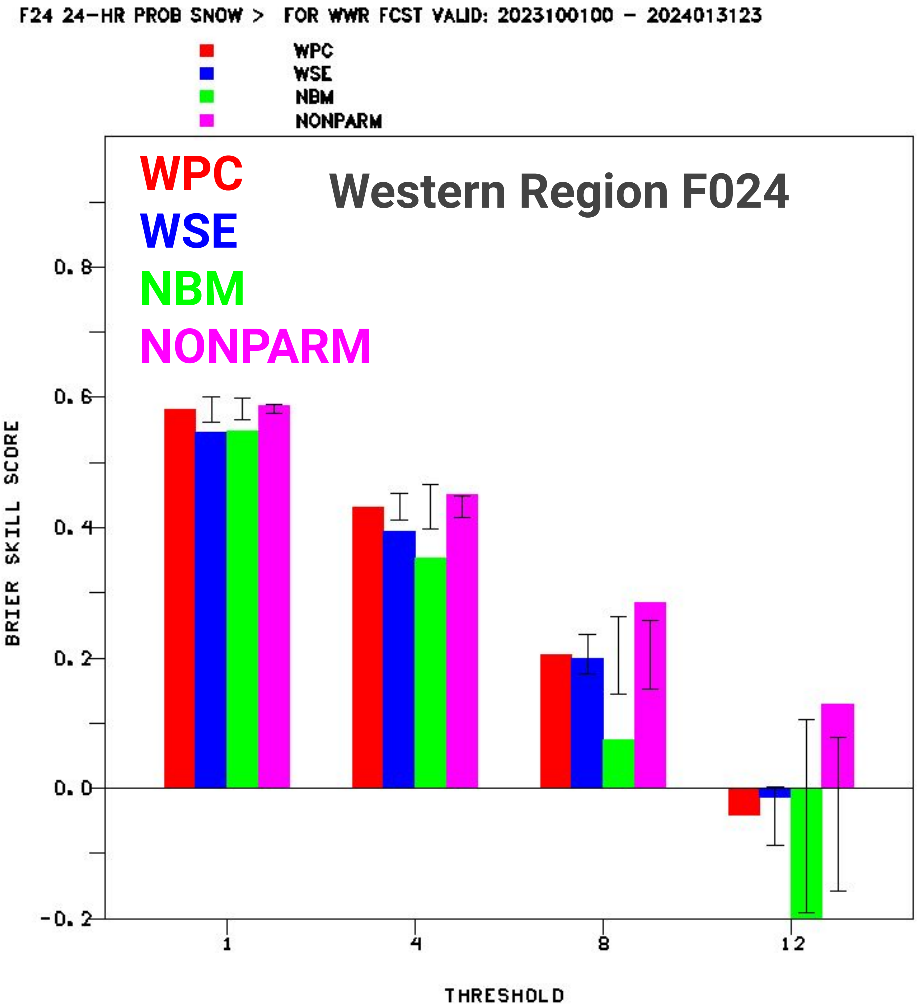
NONPARM scored higher at all thresholds with a statistically significant improvements at 4, 8, and 12 inches



Brier Skill Score
F024 Over NWS Western Region
1 Oct. 2023 - 31 Jan. 2024

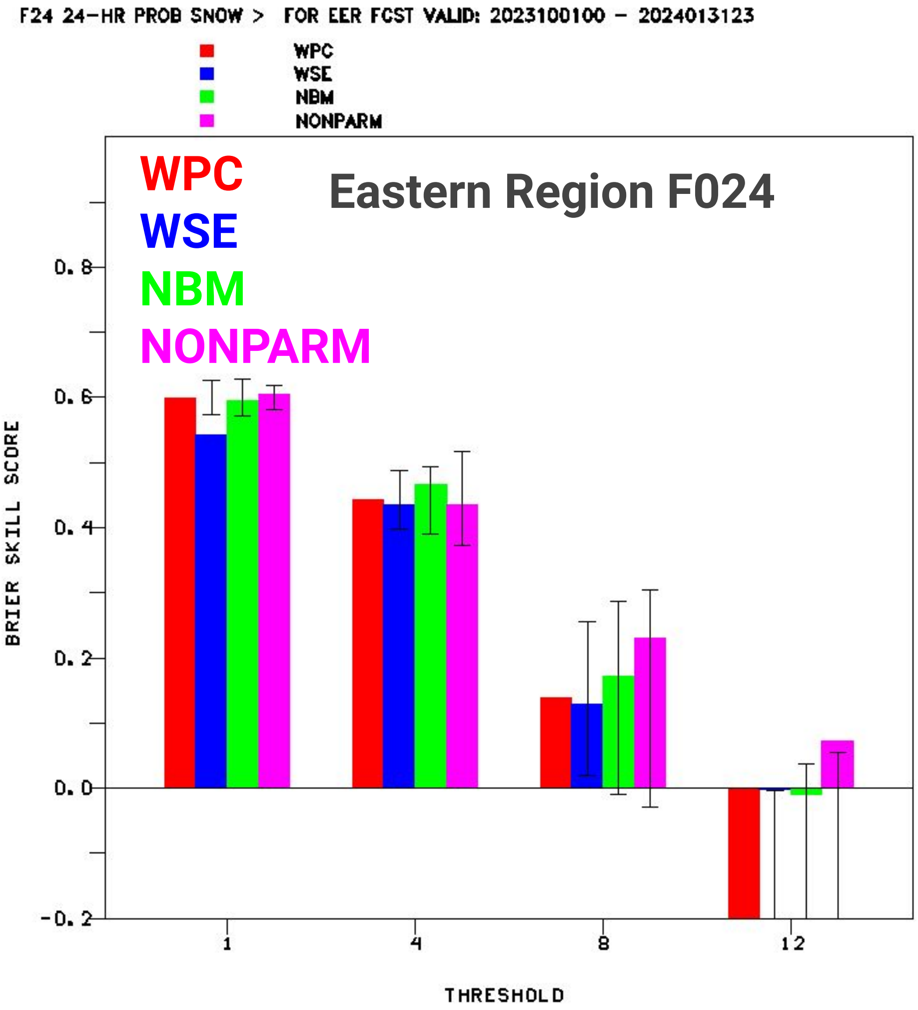
Similar results for NWS Western Region

NONPARM scored higher at all thresholds with a statistically significant improvements at 8 and 12 inches



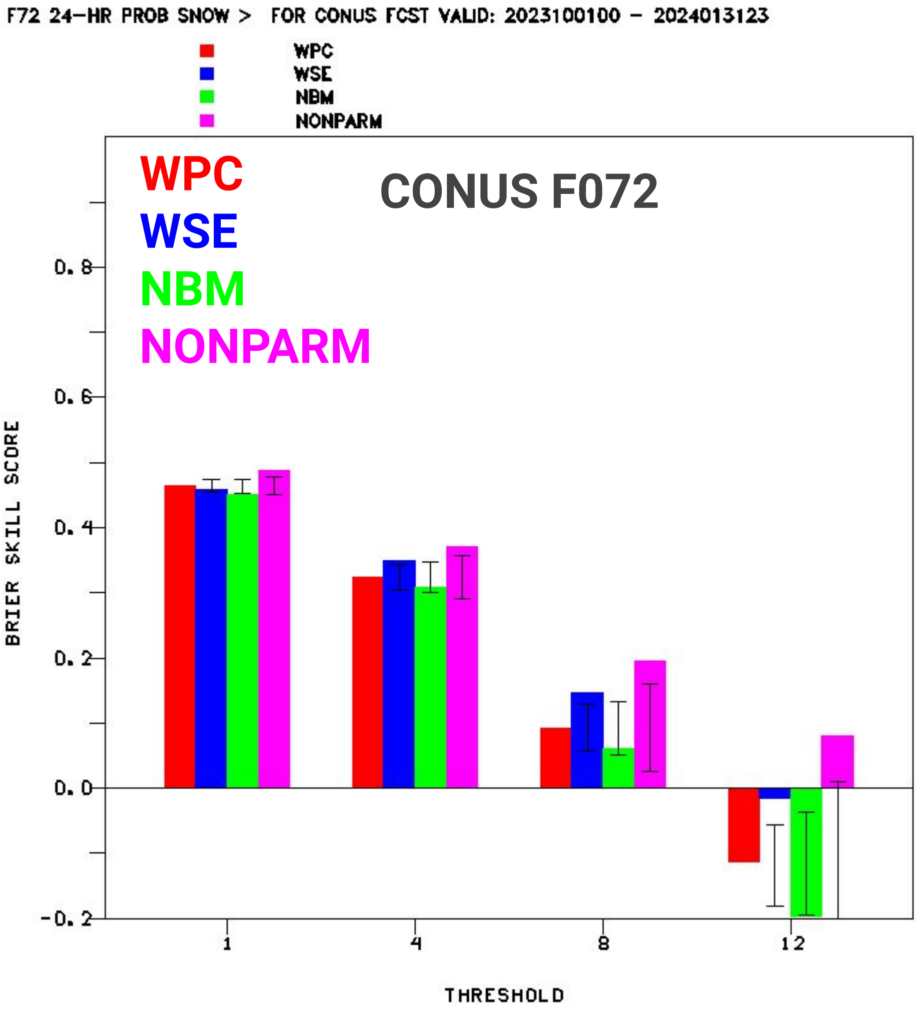
Brier Skill Score
F024 Over NWS Eastern Region
1 Oct. 2023 - 31 Jan. 2024

Results more mixed for Eastern Region except at 12 inches



Brier Skill Score
F072 Over CONUS
1 Oct. 2023 - 31 Jan. 2024

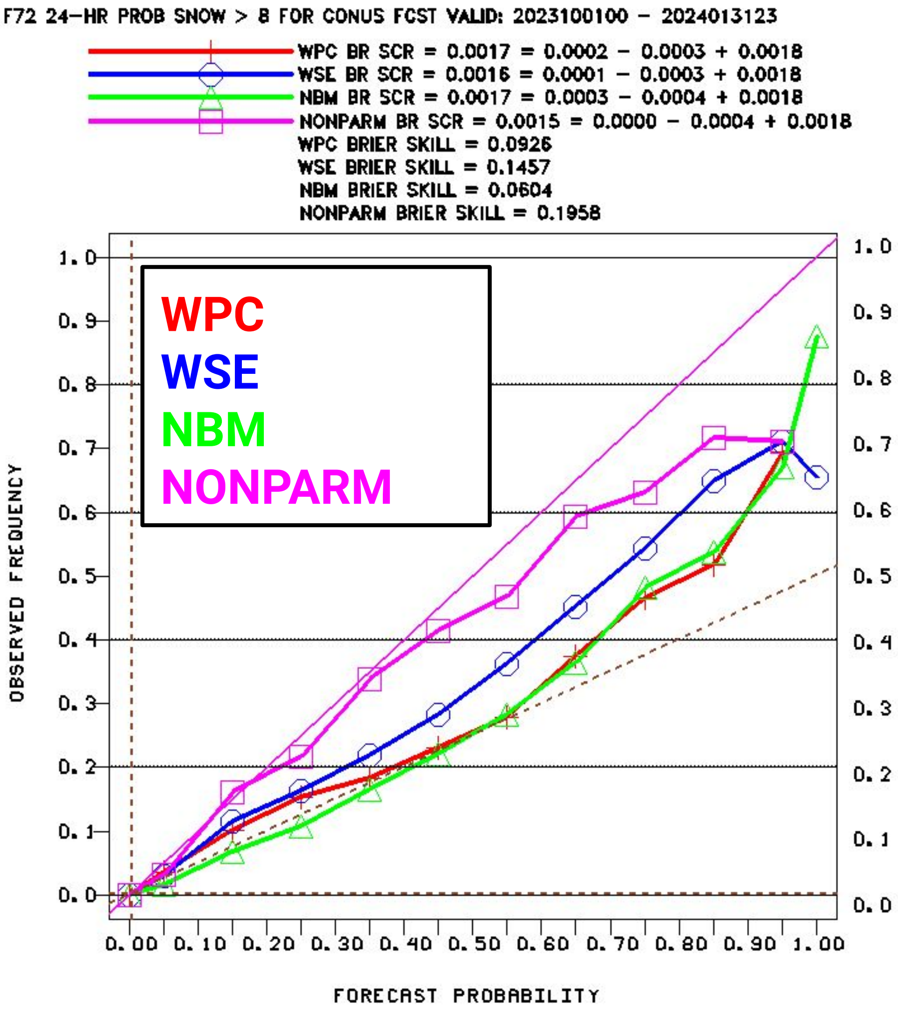
NONPARM scored higher at all thresholds with a statistically significant improvement at 1, 4, 8, and 12 inches



Reliability Prob > 8 [in]
F072 Over CONUS
1 Oct. 2023 - 31 Jan. 2024

A challenging forecast: Day 3 probability of more than 8 inches.

NONPARM reliability generally best at all forecast probability bins. Still slight overconfidence at higher probabilities.

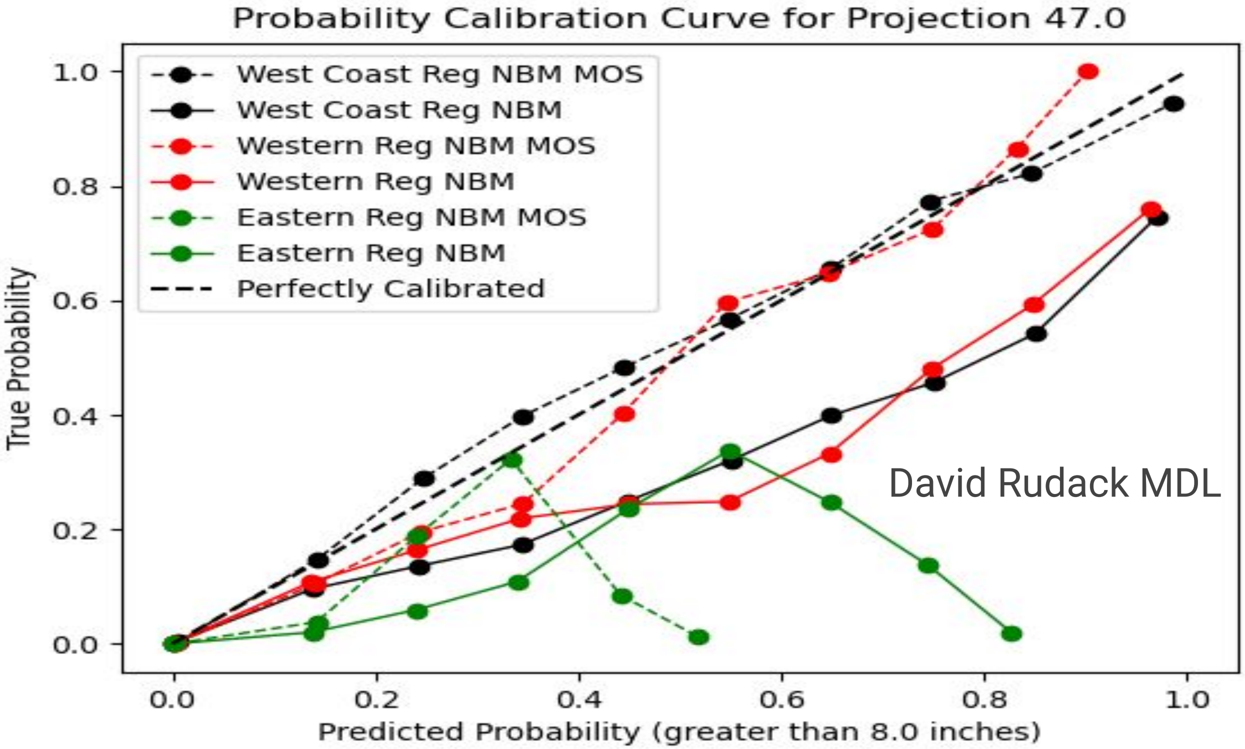


Collaboration With MDL and NBM development

David Rudack tested a similar regression-based technique using archived NBM inputs

Also showed a significant improvement in reliability compared to original NBM

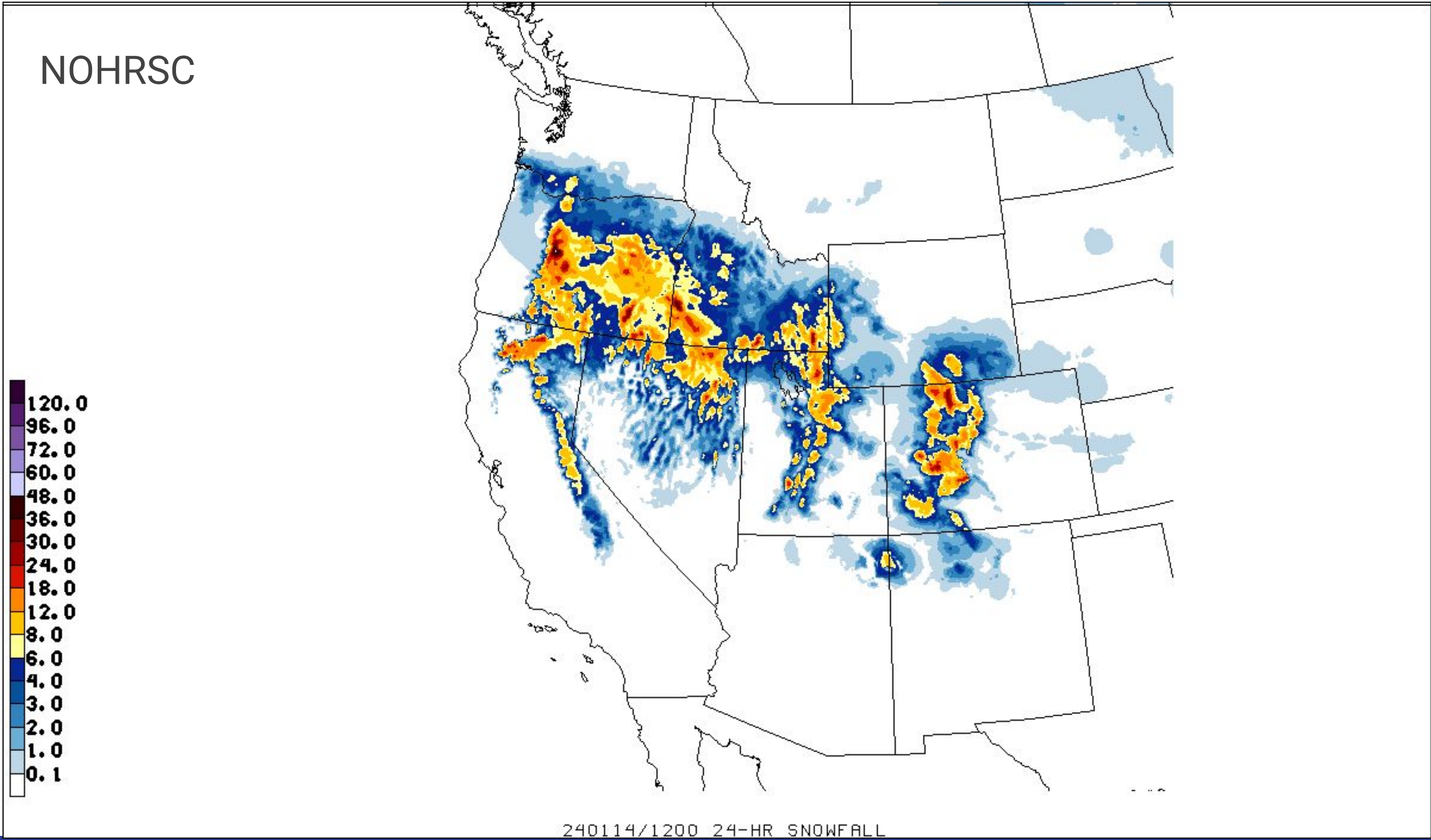
Dashed lines show result of regression calibration



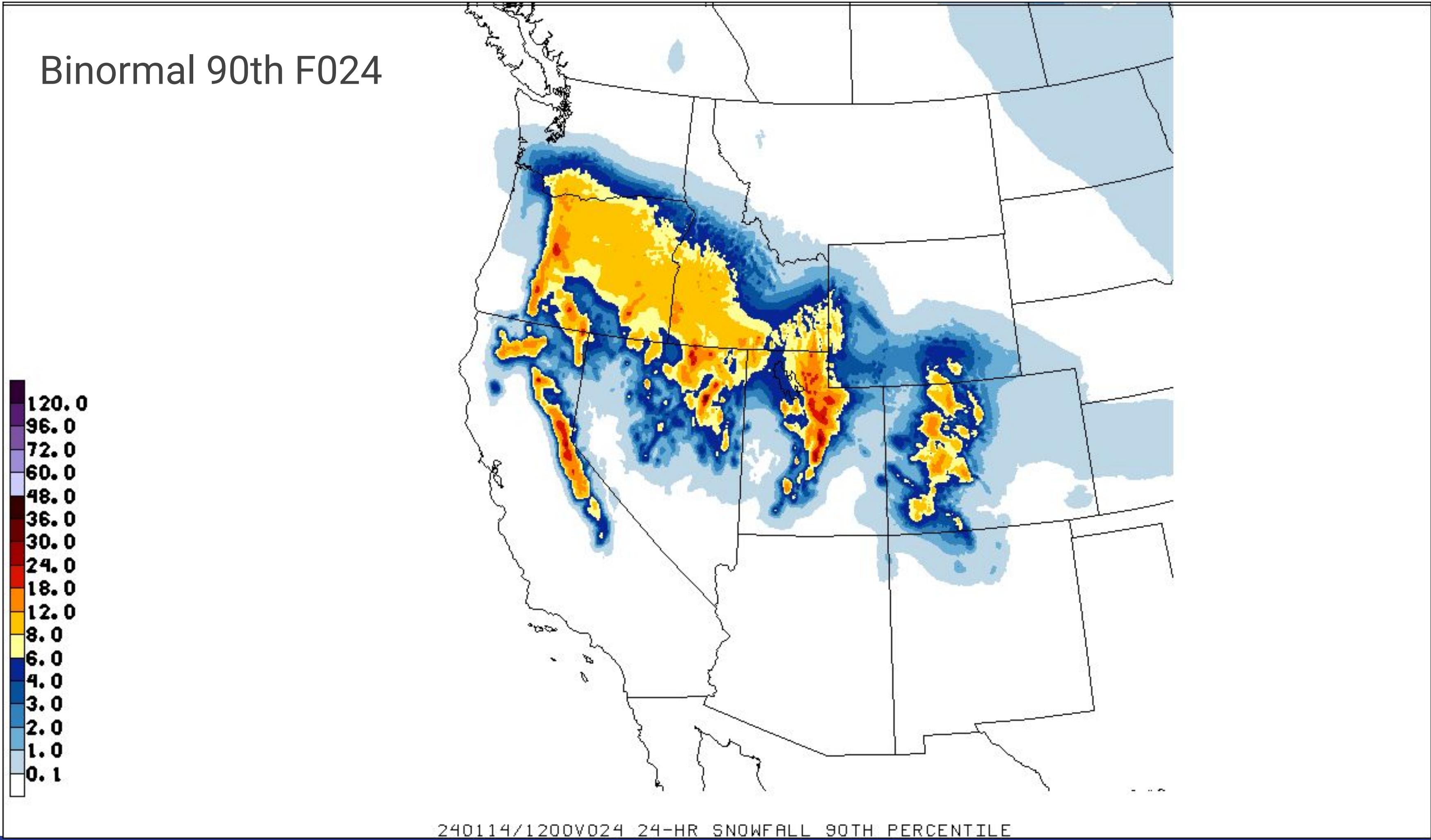
David Rudack MDL

Western US Case Study

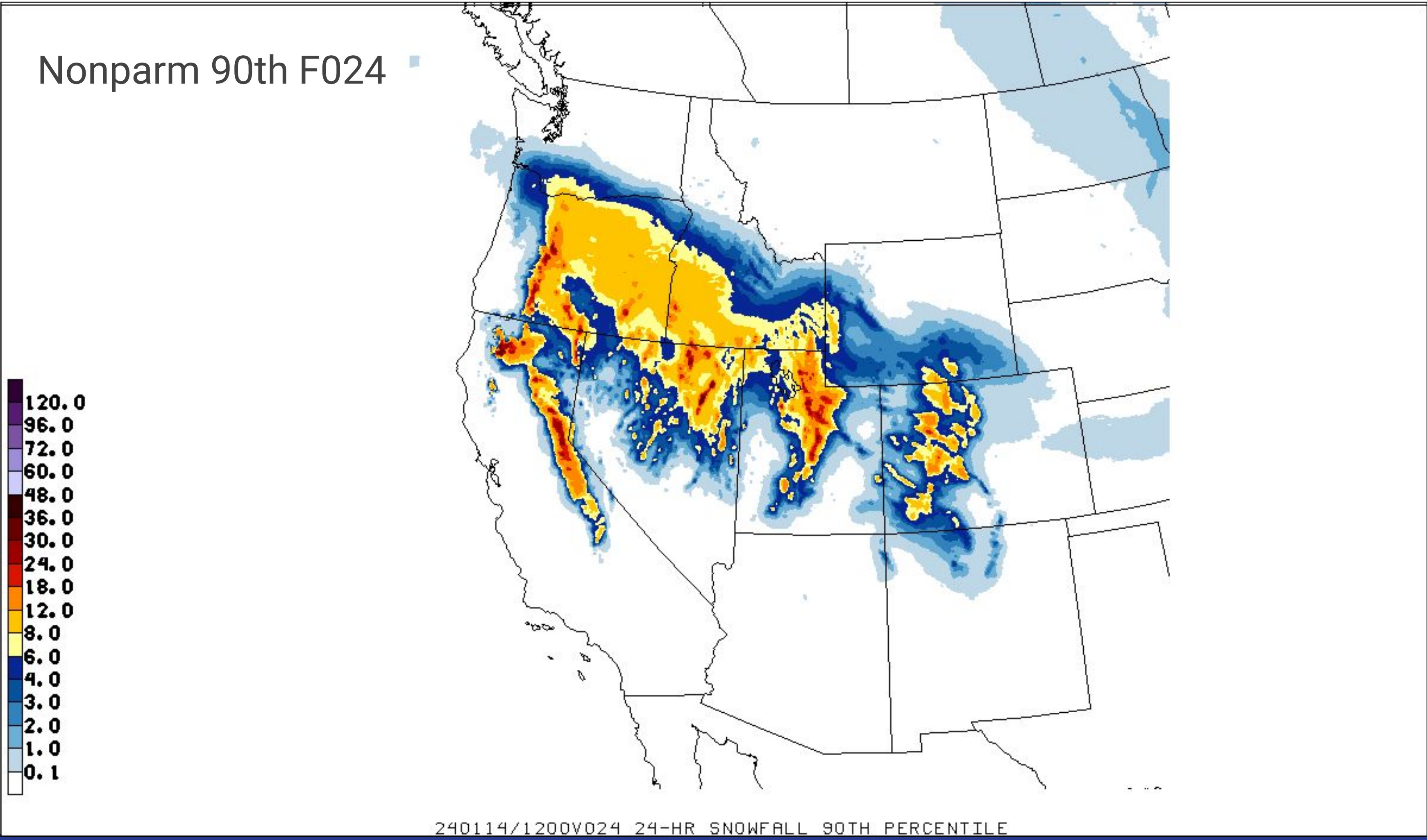
NOHRSC


120.0
96.0
72.0
60.0
48.0
36.0
30.0
24.0
18.0
12.0
8.0
6.0
4.0
3.0
2.0
1.0
0.1
240114/1200 24-HR SNOWFALL

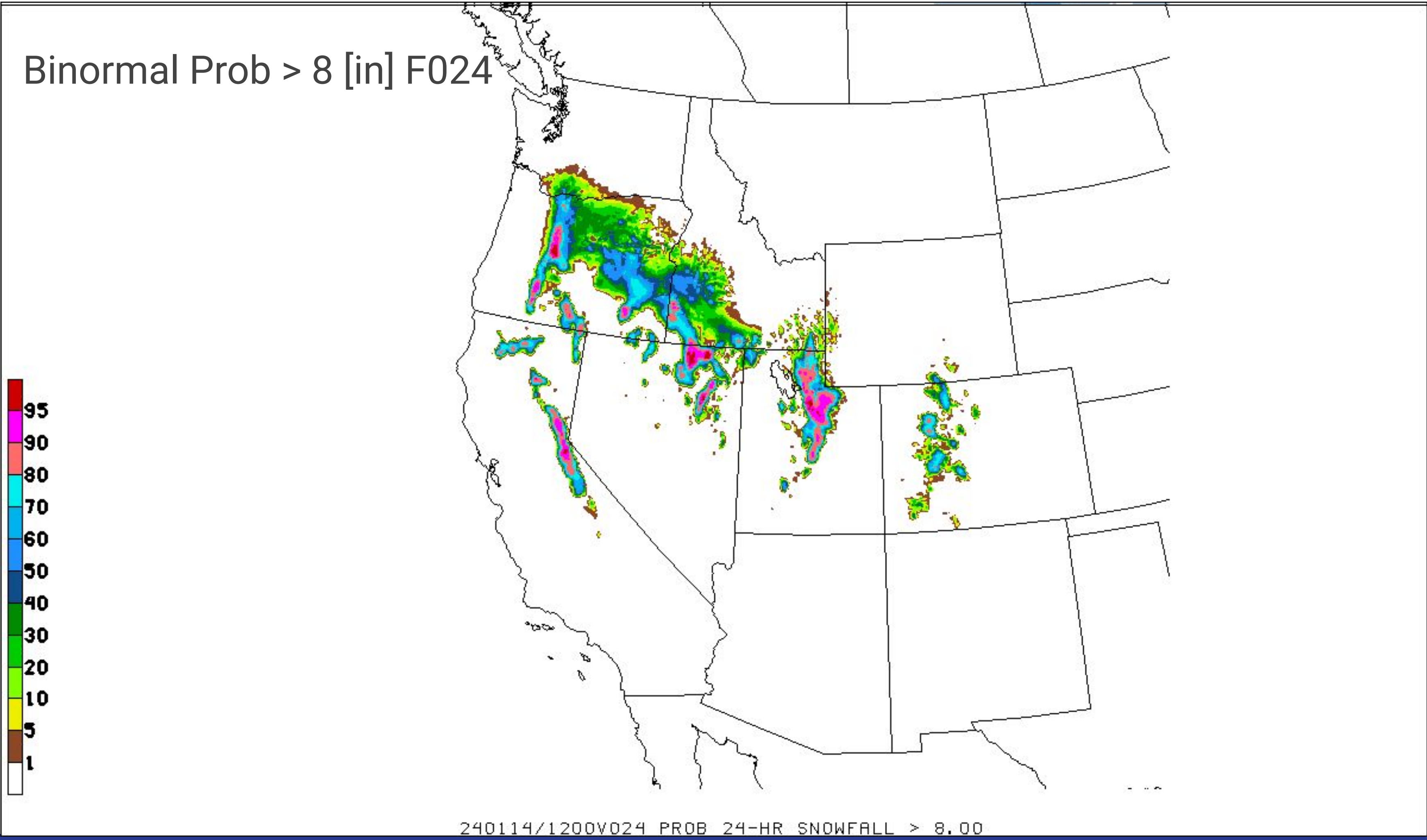
Binormal 90th F024



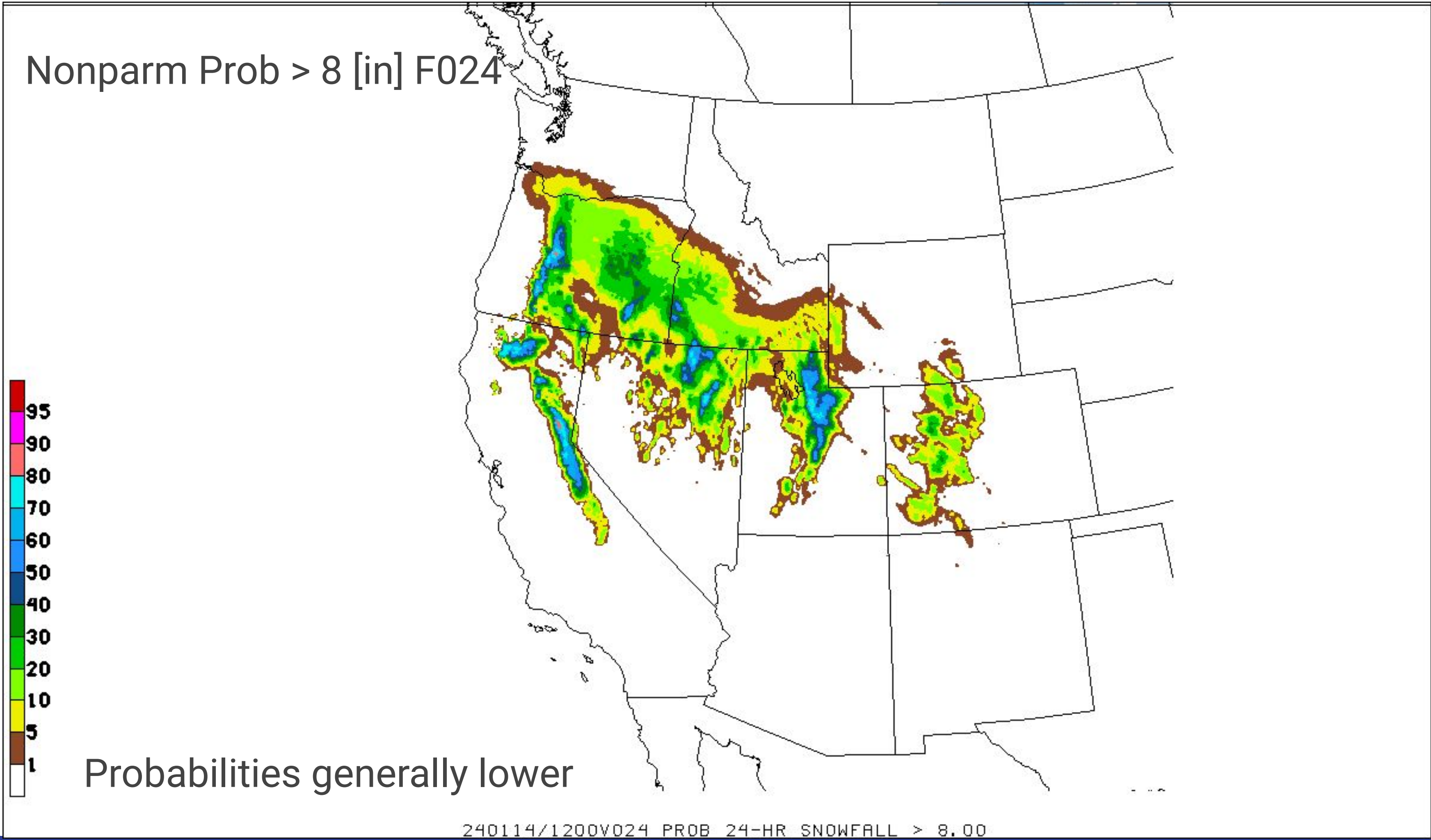
Nonparm 90th F024



Binormal Prob > 8 [in] F024

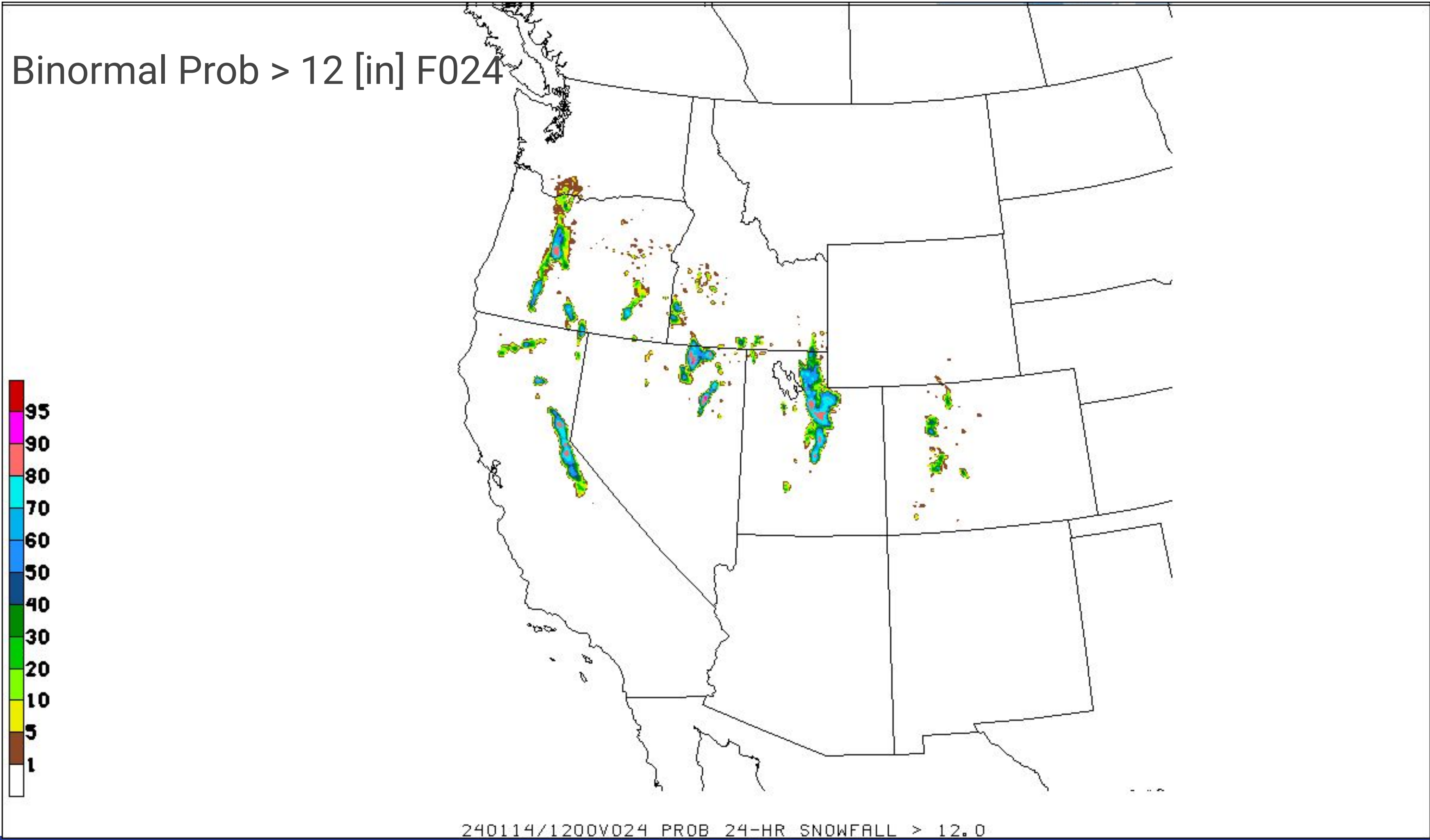


Nonparm Prob > 8 [in] F024

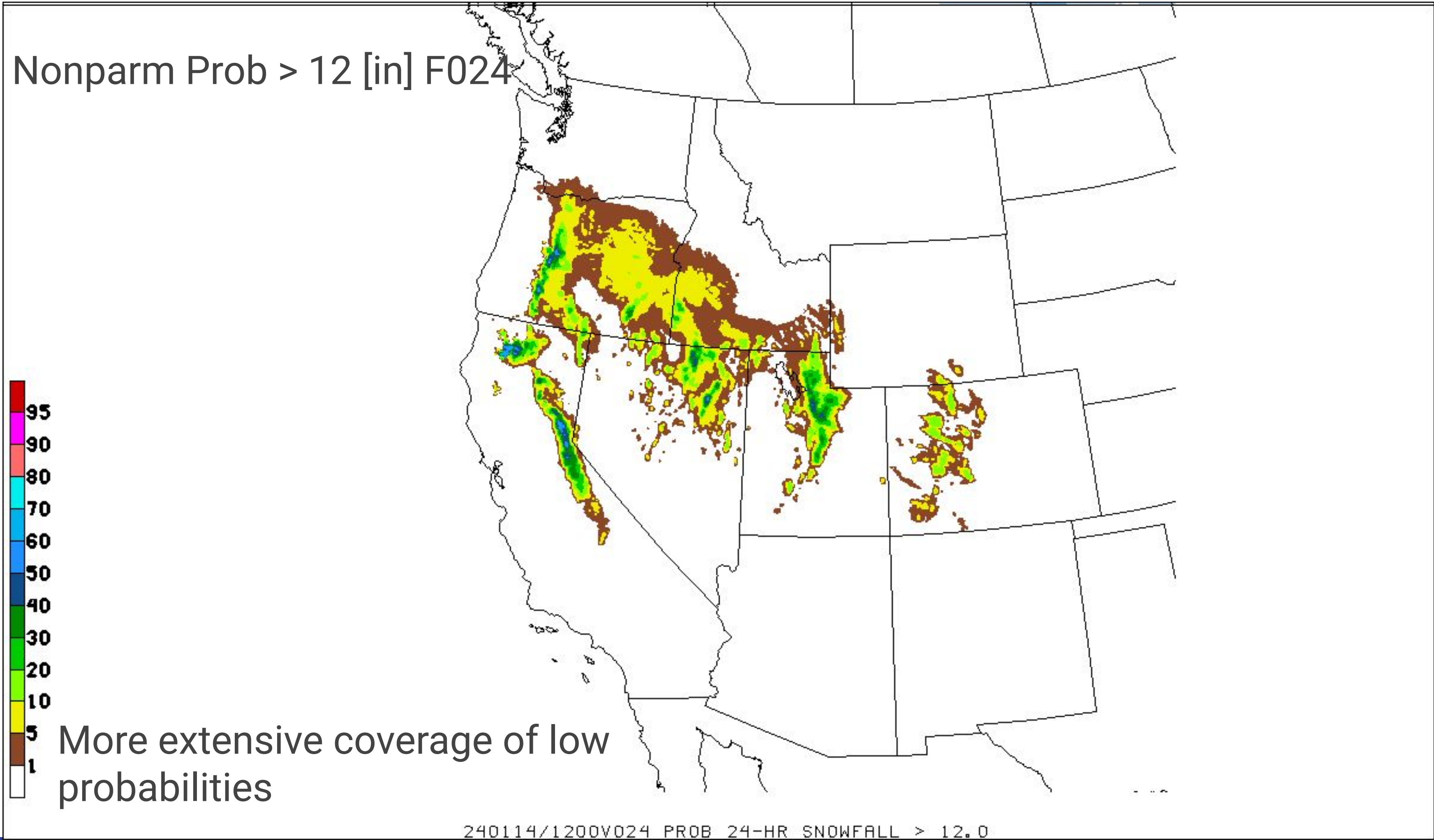


Probabilities generally lower

Binormal Prob > 12 [in] F024



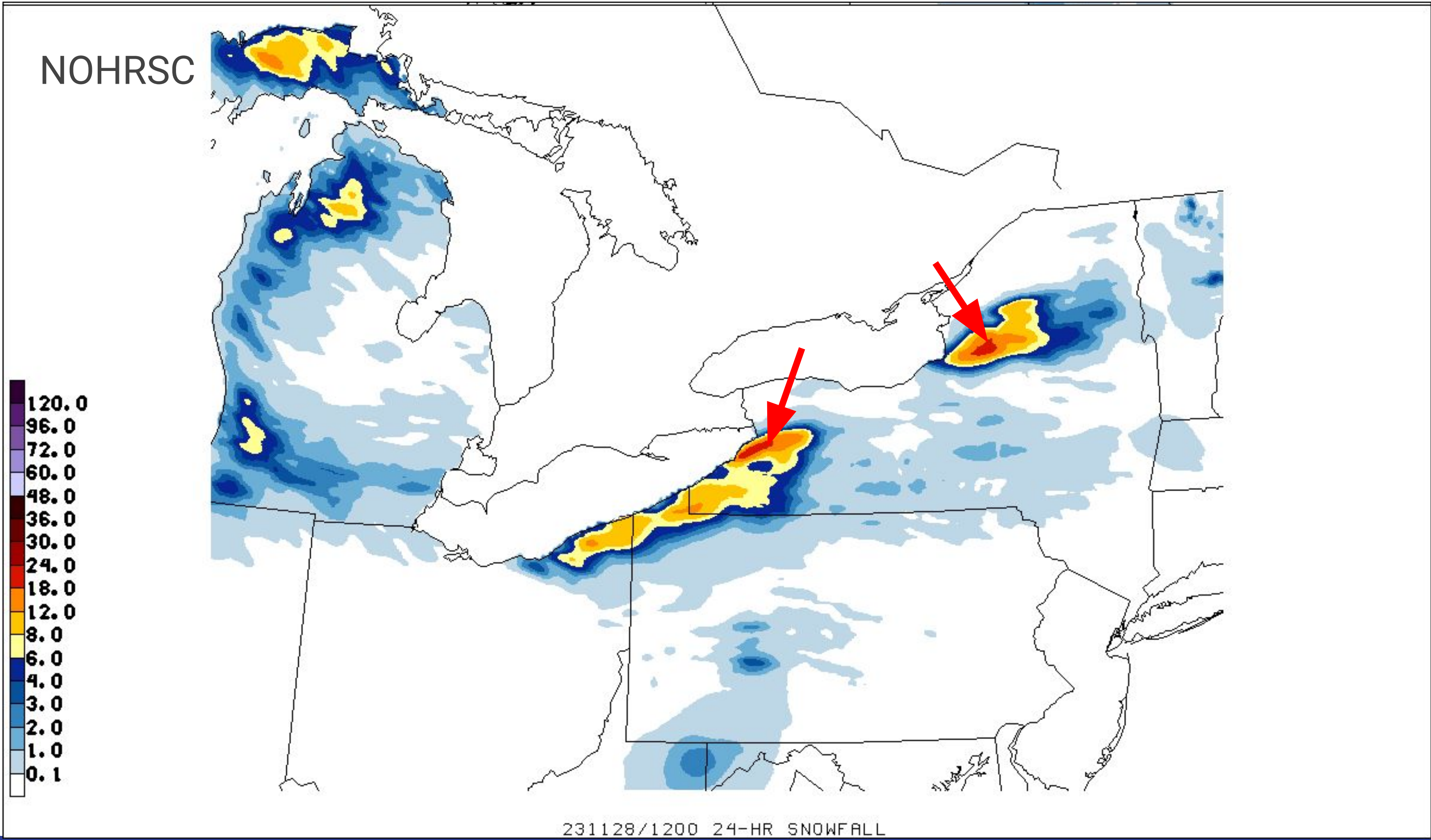
Nonparm Prob > 12 [in] F024



More extensive coverage of low probabilities

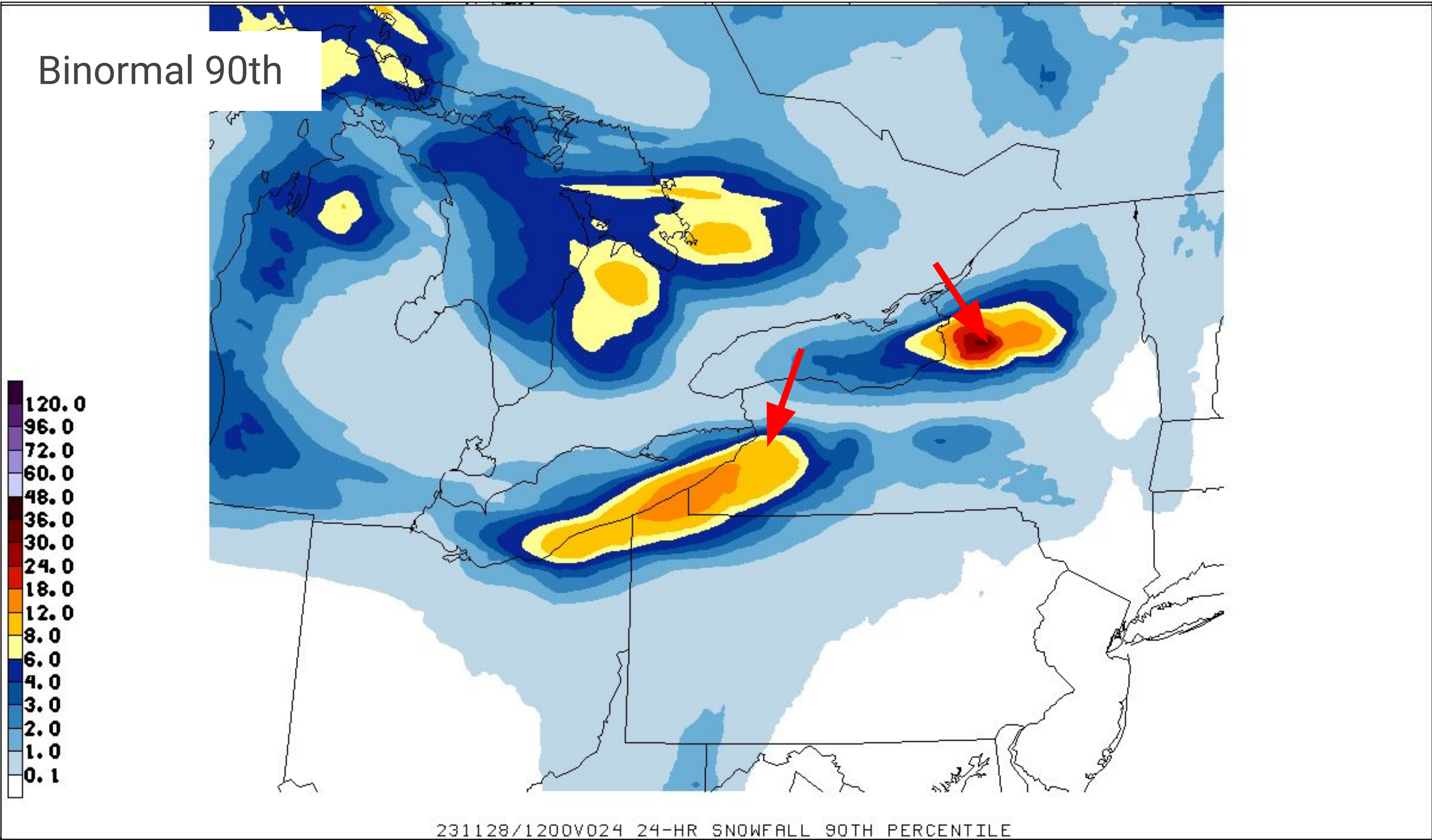
Lake Effect Snow Case Study

NOHRSC

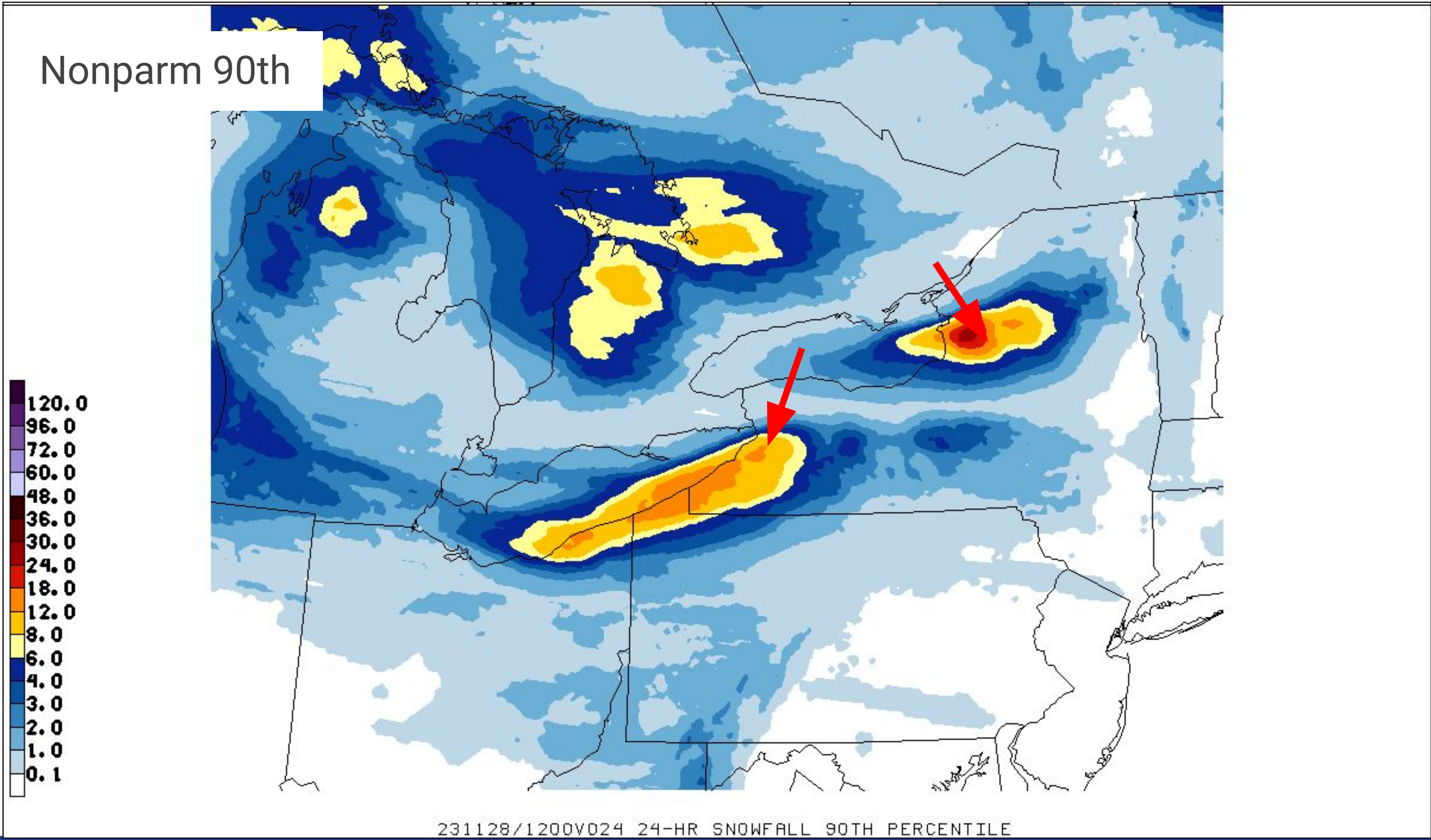


120.0
96.0
72.0
60.0
48.0
36.0
30.0
24.0
18.0
12.0
8.0
6.0
4.0
3.0
2.0
1.0
0.1
231128/1200 24-HR SNOWFALL

Binormal 90th



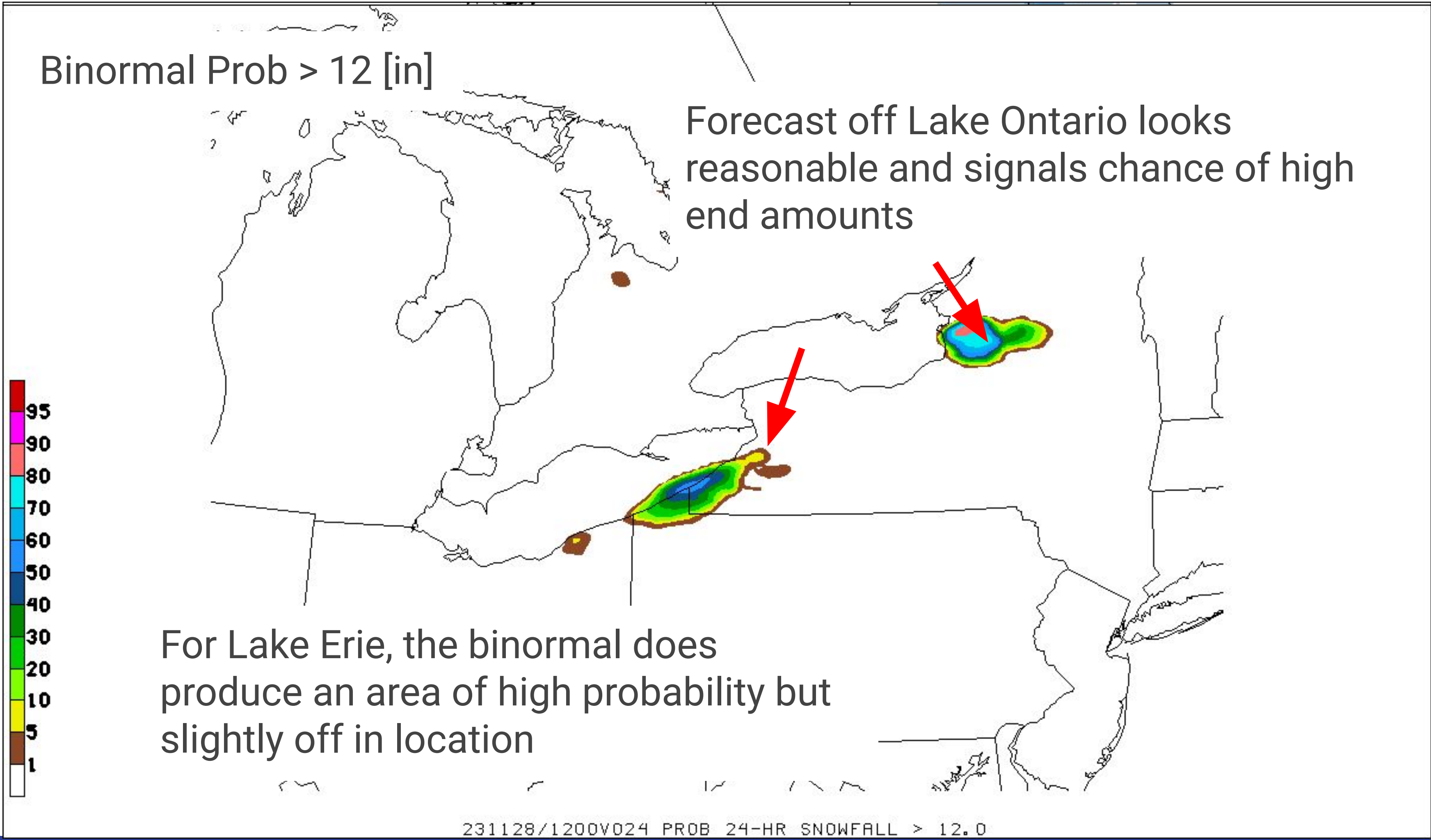
Nonparm 90th



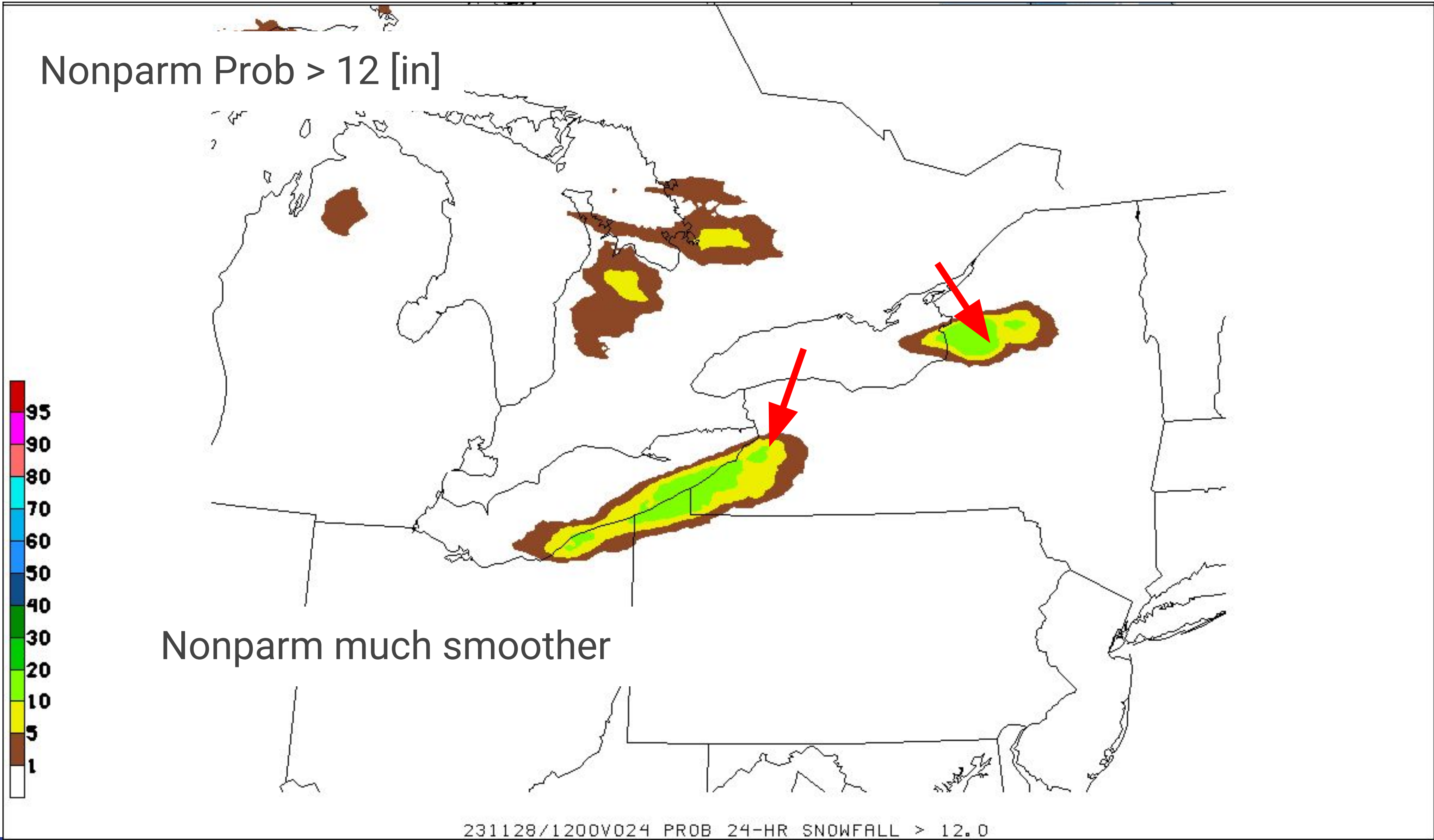
Binormal Prob > 12 [in]

Forecast off Lake Ontario looks reasonable and signals chance of high end amounts

For Lake Erie, the binormal does produce an area of high probability but slightly off in location

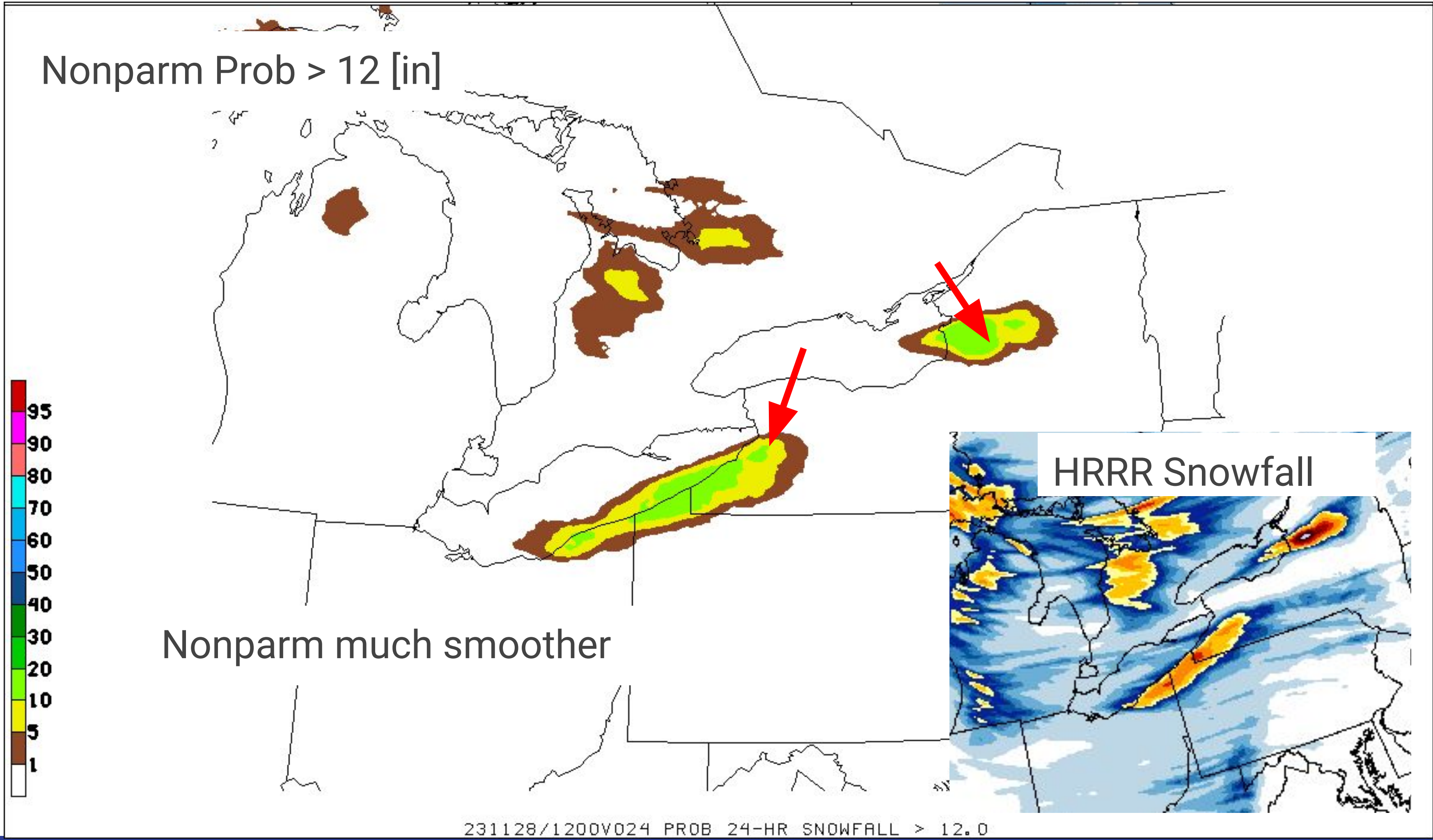


Nonparm Prob > 12 [in]



Nonparm much smoother

Nonparm Prob > 12 [in]



Nonparm much smoother

Summary

- In general the NONPARM approach scores best as measured by Brier Skill Score and reliability diagrams
- In general, if the end goal is statistical reliable forecasts then a specific calibration step is required to achieve that goal (i.e. regression or ML)
- Tradeoff between forecast sharpness and reliability
  - With calibration, would we need to consider lower probability thresholds in the watch decision process?
  - An analysis using ROC diagrams could help with understanding the POD/FAR tradeoff
- In the LES case, the binormal would likely score as overconfident, but does provide a better signal for high end amounts
- There is likely value in some type of member weighting scheme given the varying skill of individual models
- Ensemble underdispersion remains an issue. May require a much more sophisticated approach to address.

Thank you!

Nonparametric Approach Adjustment Details

5. The CDFs from steps 3 and 4 are combined as follows

   a. The following percentiles are used to approximate the final CDF (**CDF_FINAL**):
      i. [1 ,5, 10, 15, 25, 30, 40, 50, 60, 70, 75, 85, 90, 95, 99]

   b. Foreach percentile **p** in [99, 95, 90]
      i. **CDF_FINAL[p] = max(WPC_CDF[p], WSE_CAL_CDF[p])**

Take the max of the two CDFs here

   c. Foreach percentile p in [85, 75, 70, 60, 50, 40, 30, 25, 20]
      i. **CDF_FINAL[p] = WSE_CAL_CDF[p]**

Just use the calibrated WSE for these percentiles

   d. Foreach percentile p in [15, 10, 5, 1]
      i. CDF_FINAL[p] = WPC_CDF[p] where (WPC_FCST < WSE_CAL_CDF[15th])
      ii. CDF_FINAL[p] = WPC_CAL_CDF[p] where (WPC_FCST >= WSE_CAL_CDF[15th])

Where WPC is less than the calibrated WSE 15th use the quantile regression result. Otherwise use the calibrated WSE.

Nonparametric Approach Steps: Part 3

6. Compute the desired probability of exceedance grids from **CDF_FINAL**