Application of Multi-Dimensional Stratification in Forecast Verification

Mike Baldwin
CIWRO/SPC

image from 14 Nov 2022 I-40 OK/TX border
https://its.txdot.gov/its/District/CHS

Transportation Applications Team

Heather Reeves, PhD
Andrew Rosenow, PhD
Michael Baldwin, PhD
Daniel Tripp, MS
Adam Werkema, MS
Jilliann Dufort, MS

NOAA/National Severe Storms Laboratory

OU/CIWRO *(Cooperative Institute for Severe and High-Impact Weather Research and Operations)*

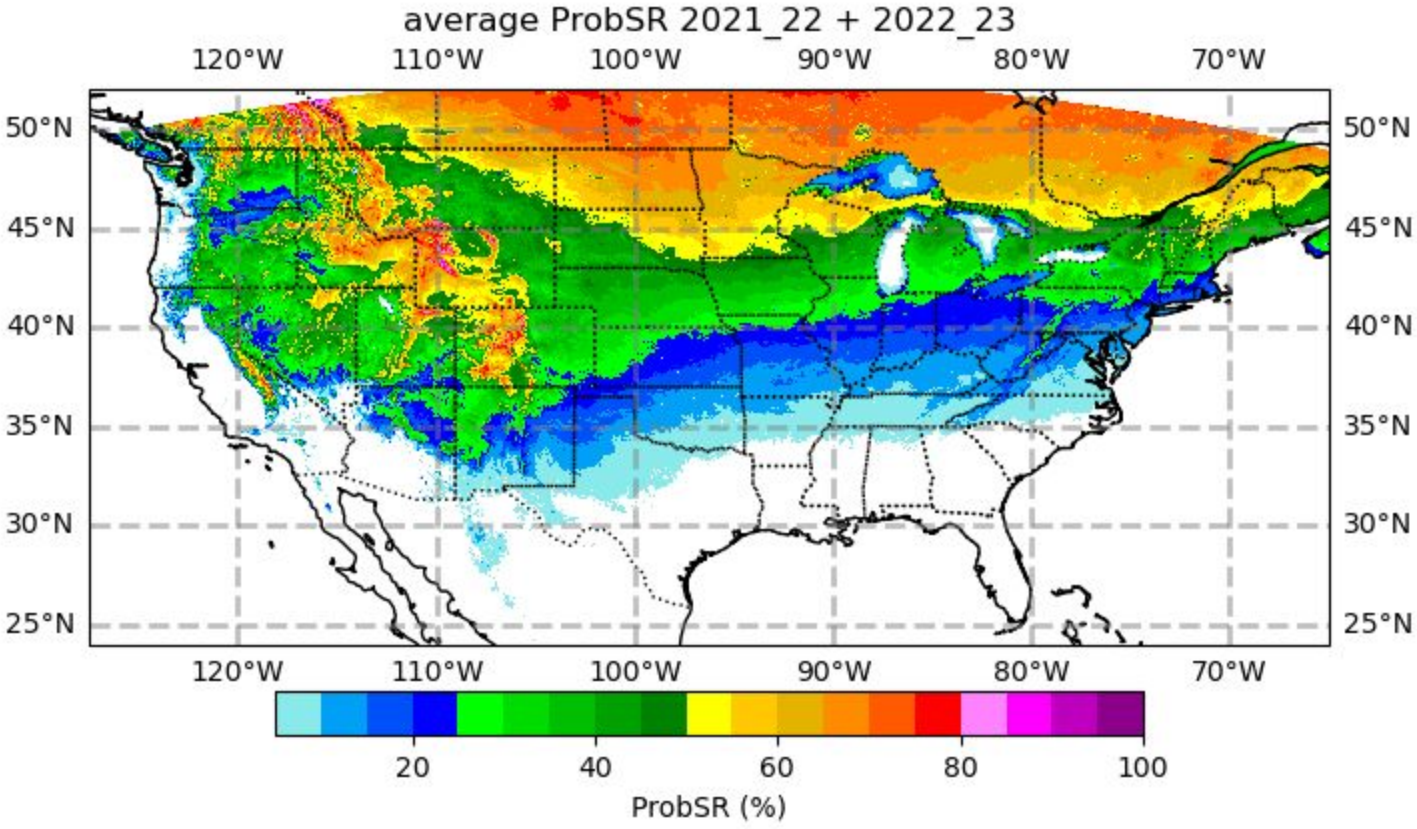
What we do:

develop decision-support tools for the DOT and NWS

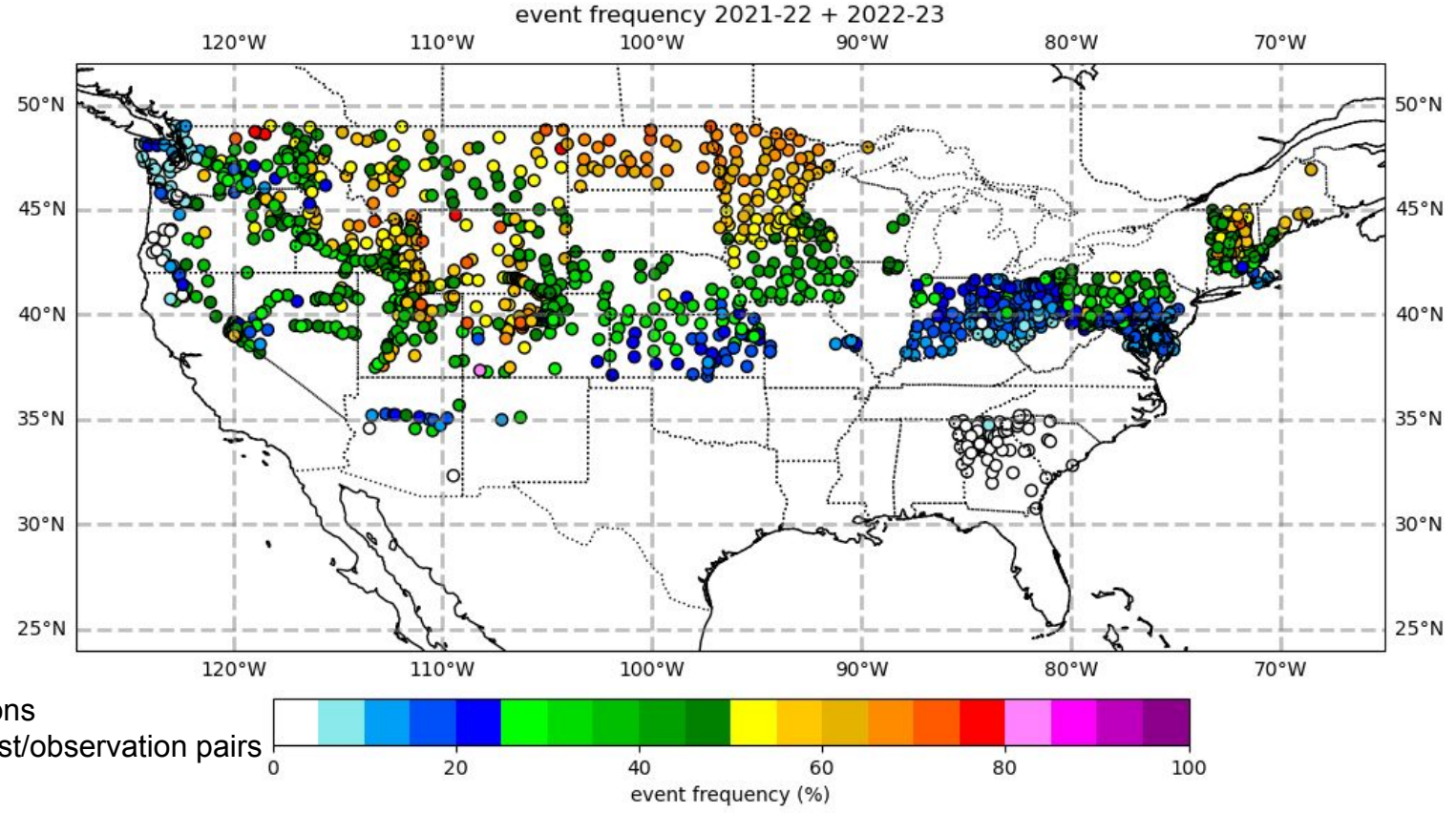
Project themes:

road weather, aviation weather, marine weather

ProbSR: Probability of Sub-freezing Roads

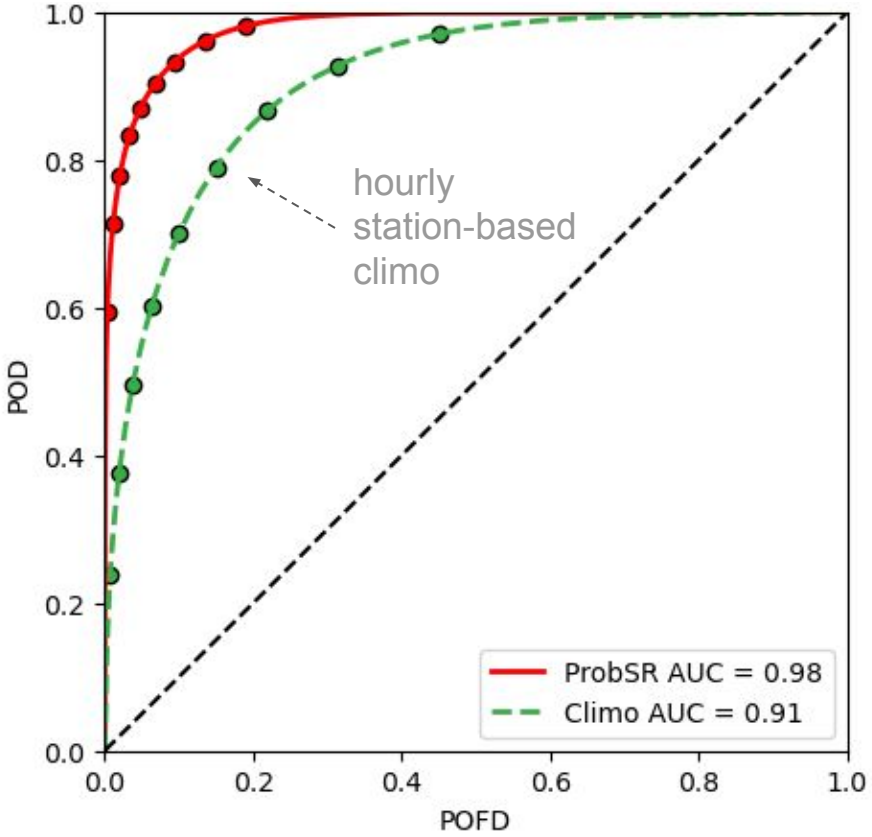


RWIS: sensors to measure road surface temperature

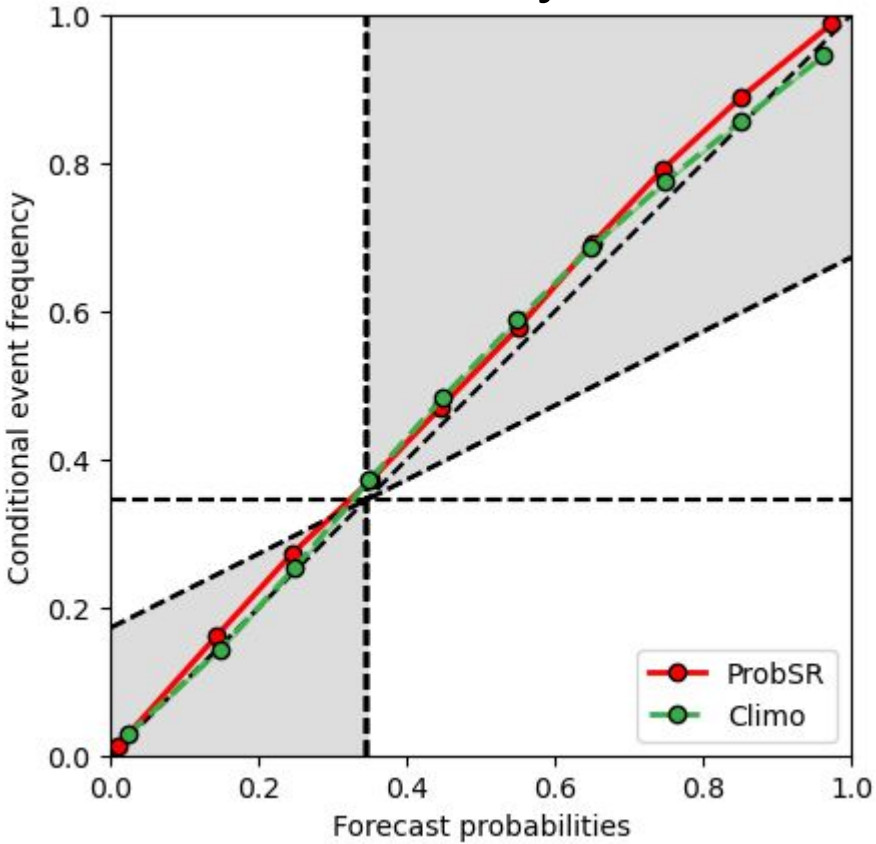


~1500 stations
9.3M forecast/observation pairs

ROC



reliability

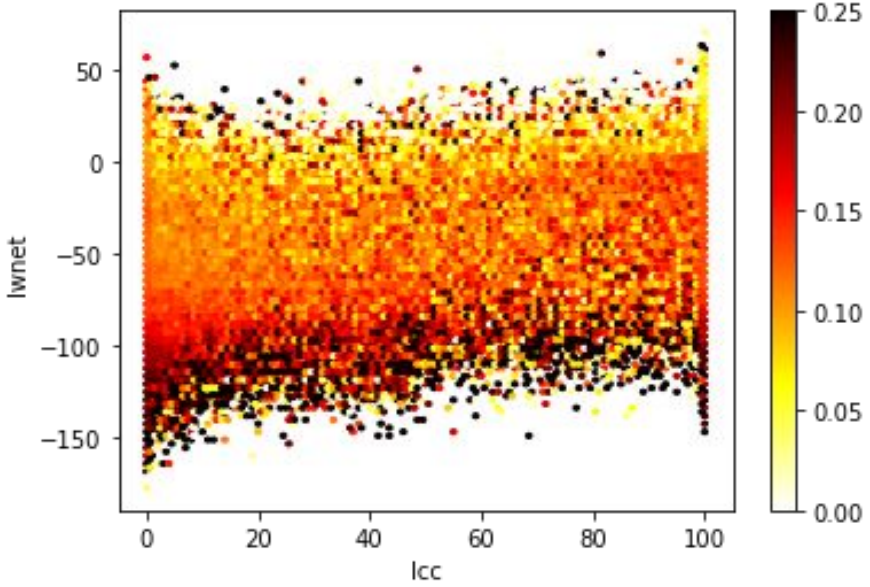


pandas

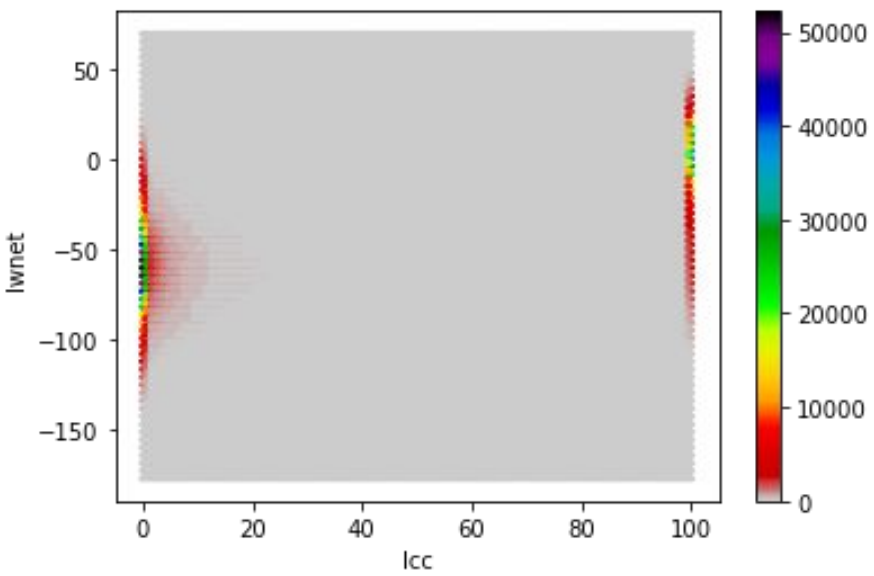
low cloud % (lcc) vs net longwave radiation (lwnet)

forecasts appear to get better as lwnet increases

mean squared error



number of events



stratification

$$bias = \sum_{k=1}^{m} \frac{n_k}{N} bias_k = \sum_{k=1}^{m} \frac{n_k}{N} (\bar{f}_k - \bar{x}_k)$$

split up the verification dataset into subsets

MSE, bias (and other metrics) can be written in terms of partial sums

$$MSE_k = \frac{1}{n_k} \sum_{j=1}^{n_k} (f_{kj} - x_{kj})^2$$

MSE for k[th] subset

$$MSE = \sum_{k=1}^{m} \frac{n_k}{N} MSE_k = \frac{1}{N} \sum_{k=1}^{m} n_k \left[\frac{1}{n_k} \sum_{j=1}^{n_k} (f_{kj} - x_{kj})^2 \right]$$

there are m subsets,
overall MSE is simple
weighted-sum of MSE_k

can subset the dataset anyway we'd like:
easy/difficult forecasts, by hour, by region,
using external variables (precip, radiation, cloud, …)

$$SS = \frac{MSE_{ref} - MSE}{MSE_{ref}}$$

$$SS_k = \frac{MSE_{ref_k} - MSE_k}{MSE_{ref_k}}$$

$$SS = \sum_{k=1}^{m} \frac{n_k}{N} \frac{MSE_{ref_k}}{MSE_{ref}} SS_k$$

relative frequency of k^{th} subset

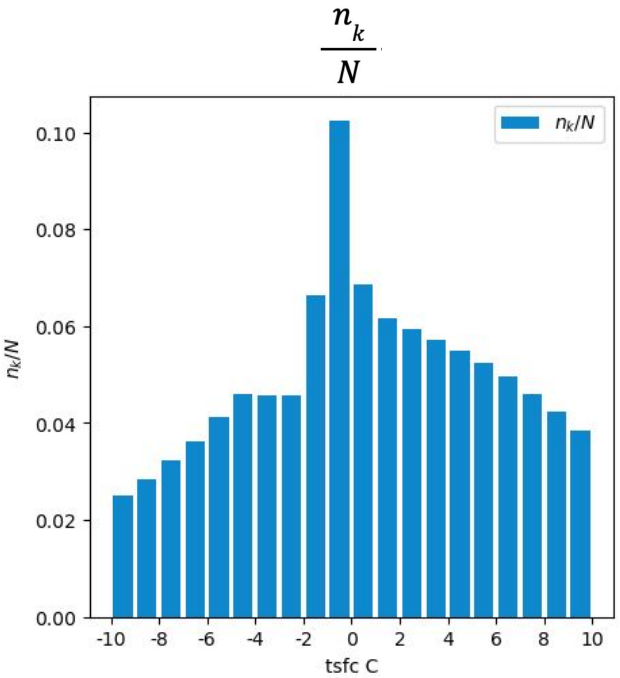
normalized MSE_{ref} of k^{th} subset

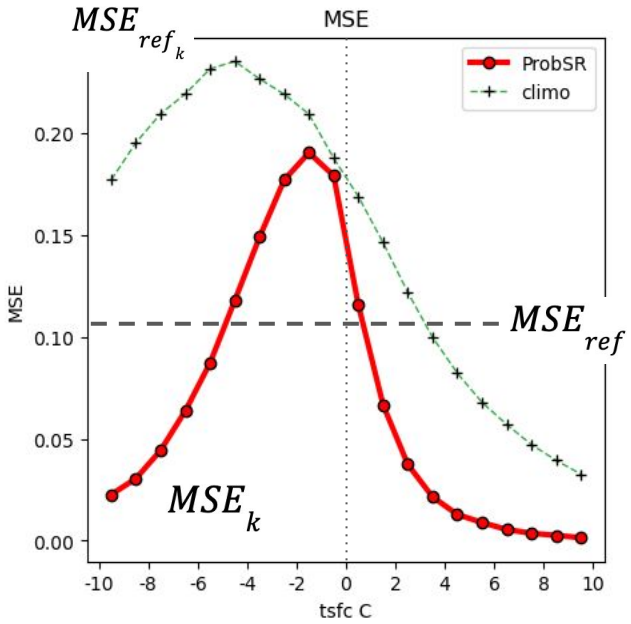
measure of forecast “challenge”

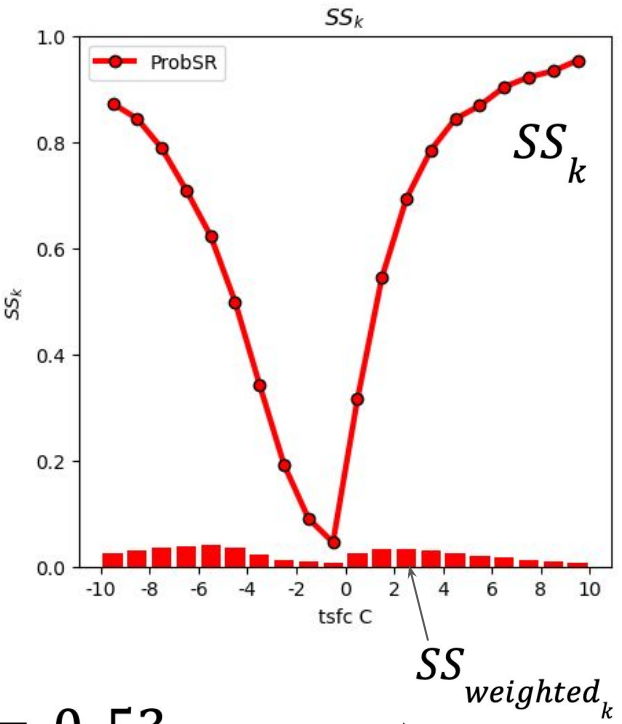
to determine the contribution of k^{th} subset to the *overall* skill score, need to multiply SS_k by both relative frequency and normalized reference MSE

- more weight to more challenging subsets
- more weight to more frequent subsets

stratify by HRRR T_{sfc}







$$SS = \sum_{k=1}^{m} \frac{n_k}{N} \frac{MSE_{ref_k}}{MSE_{ref}} SS_k = \sum_{k=1}^{m} SS_{weighted_k} = 0.53$$

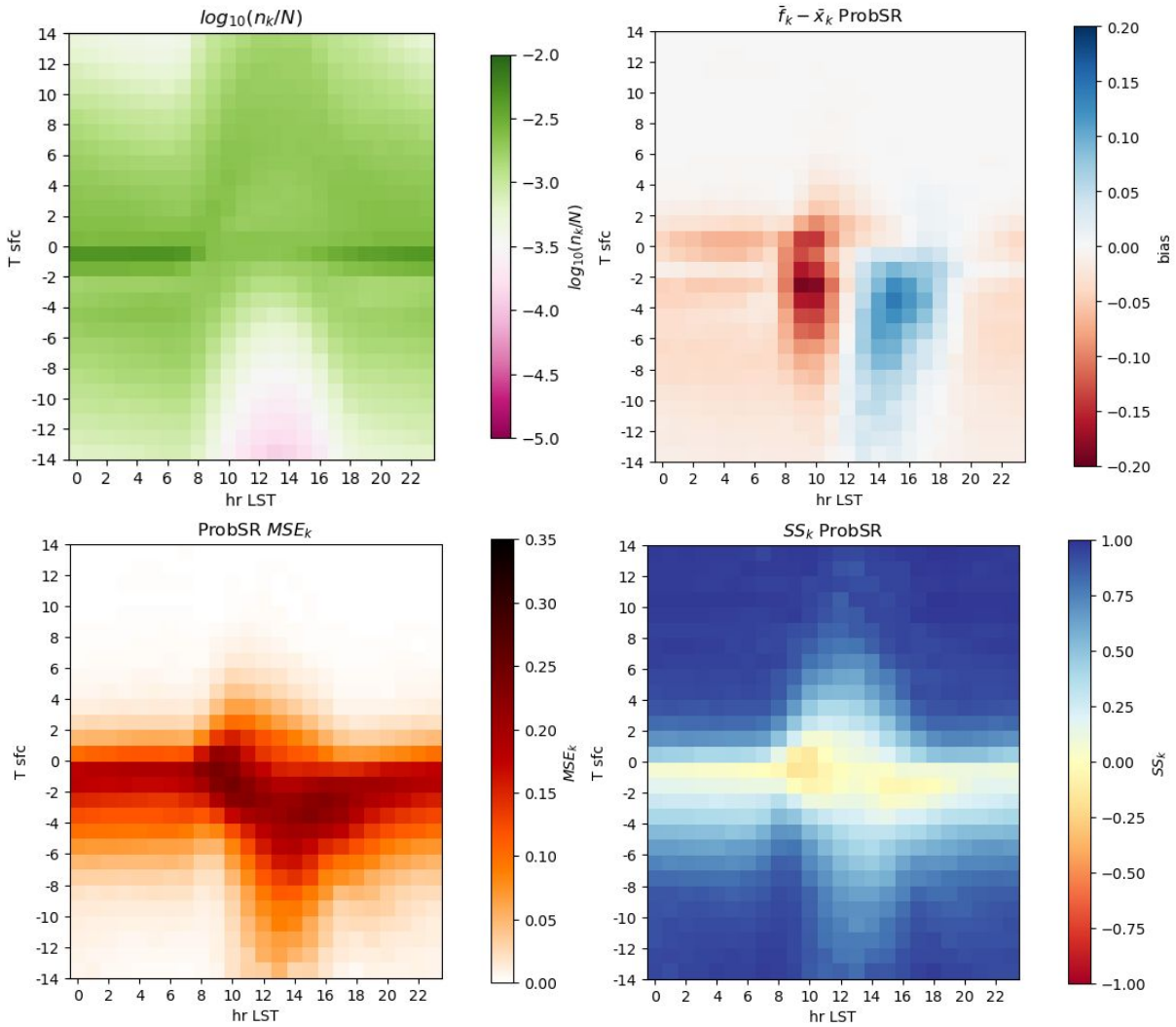
stratification across multiple dimensions

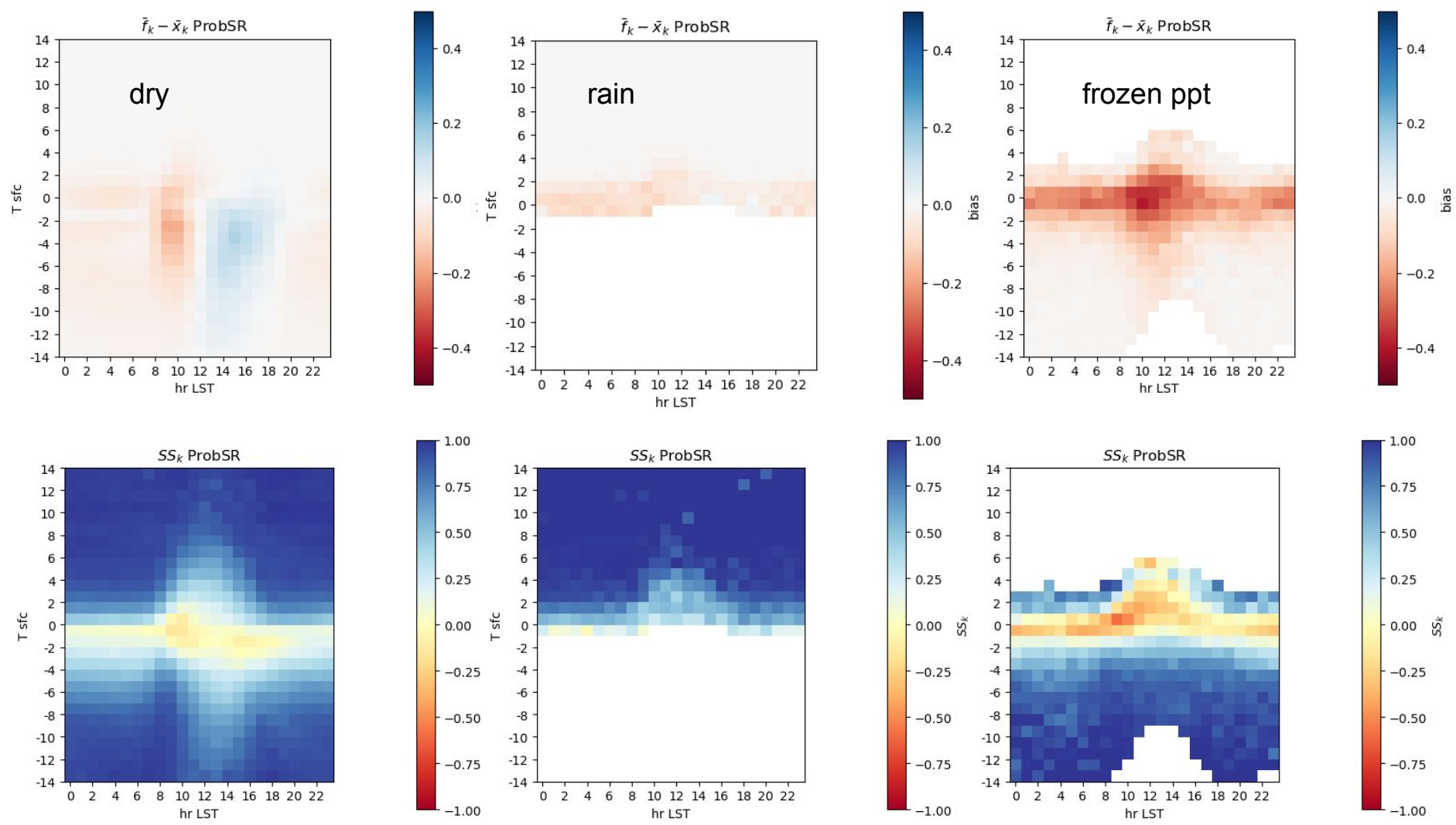
2-D stratification by

HRRR T_{sfc}

and

hour of the day (LST)




$\bar{f}_k - \bar{x}_k$ ProbSR
dry
rain
frozen ppt
T sfc
hr LST
bias
SS_k ProbSR
SS_k

summary - future work

many dimensions/variables could be used for stratification

challenge: determine relevant “axes” (covariates) that allow for

- discovery of conditional biases/errors
- focused forecast system improvements
- boosted confidence in situations with greatest skill

reduce dimensionality (PCA/cluster analysis/mixture models)

could be an opportunity for AI/ML