Statistics for HMT-West ensemble precipitation forecasts during the winter 2009-2010

Huiling Yuan[1, 2], Isidora Jankov[1, 3], Steve Albers[1, 3], Ed Tollerud[1], Tara Jensen[4], John Halley Gotway[4], Paul Schultz[1], Linda Wharton[1], Zoltan Toth[1], and Curtis Alexander[1, 2]

1. NOAA/ESRL/Global Systems Division
2. CIRES, University of Colorado, Boulder, CO
3. CIRA, Colorado State University, Fort Collins, CO
4. NCAR/Developmental Testbed Center (DTC)

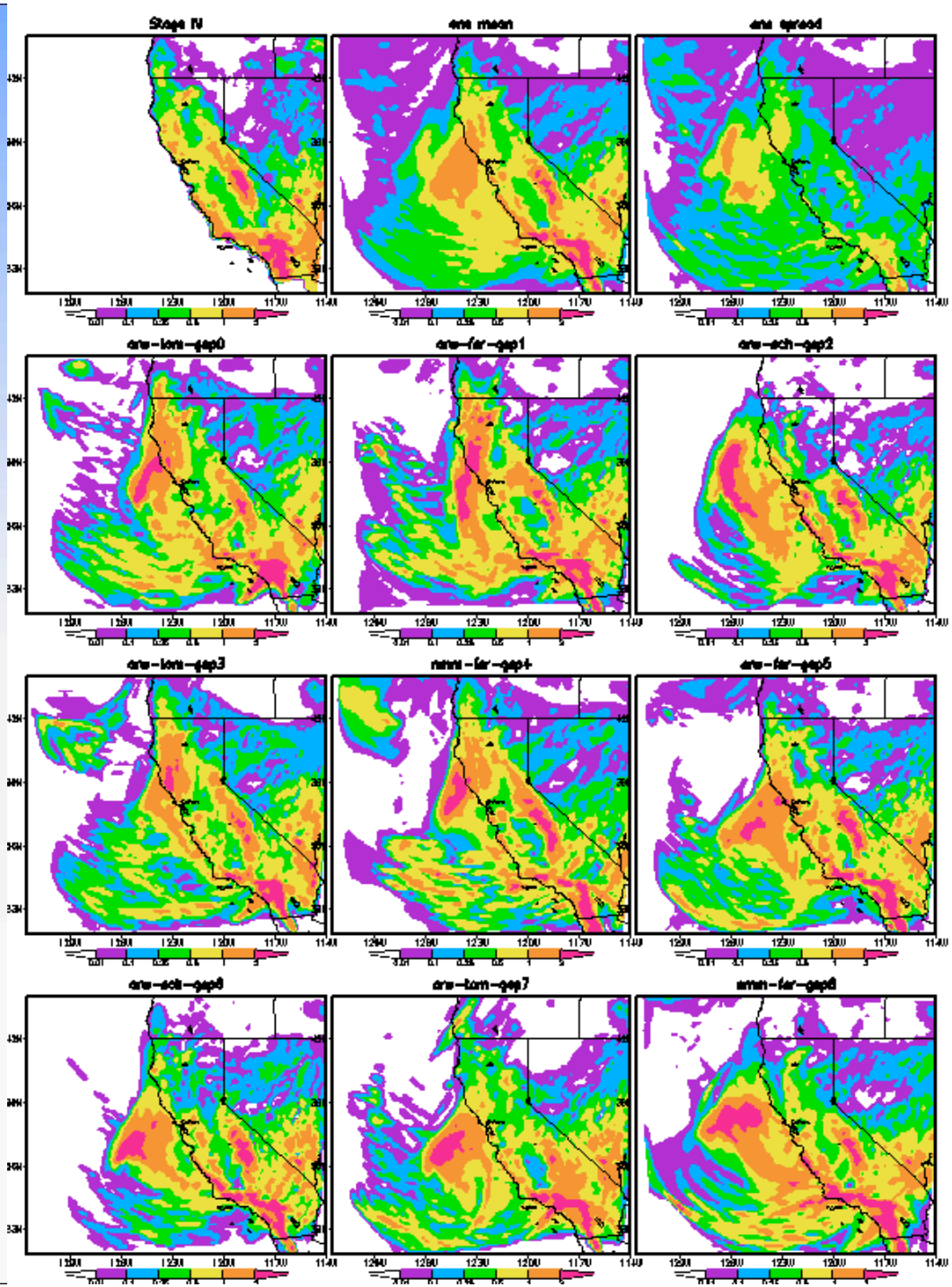
Outline

- Examples of Quantitative Precipitation Forecasts (QPF) and Probabilistic QPF (PQPF)
- Statistics for Dec 2009 – April 2010:
- Deterministic scores (ETS)
- Probabilistic verification (reliability, BSS)
- Future work

QPF

Example of 24-h QPF
9-km resolution

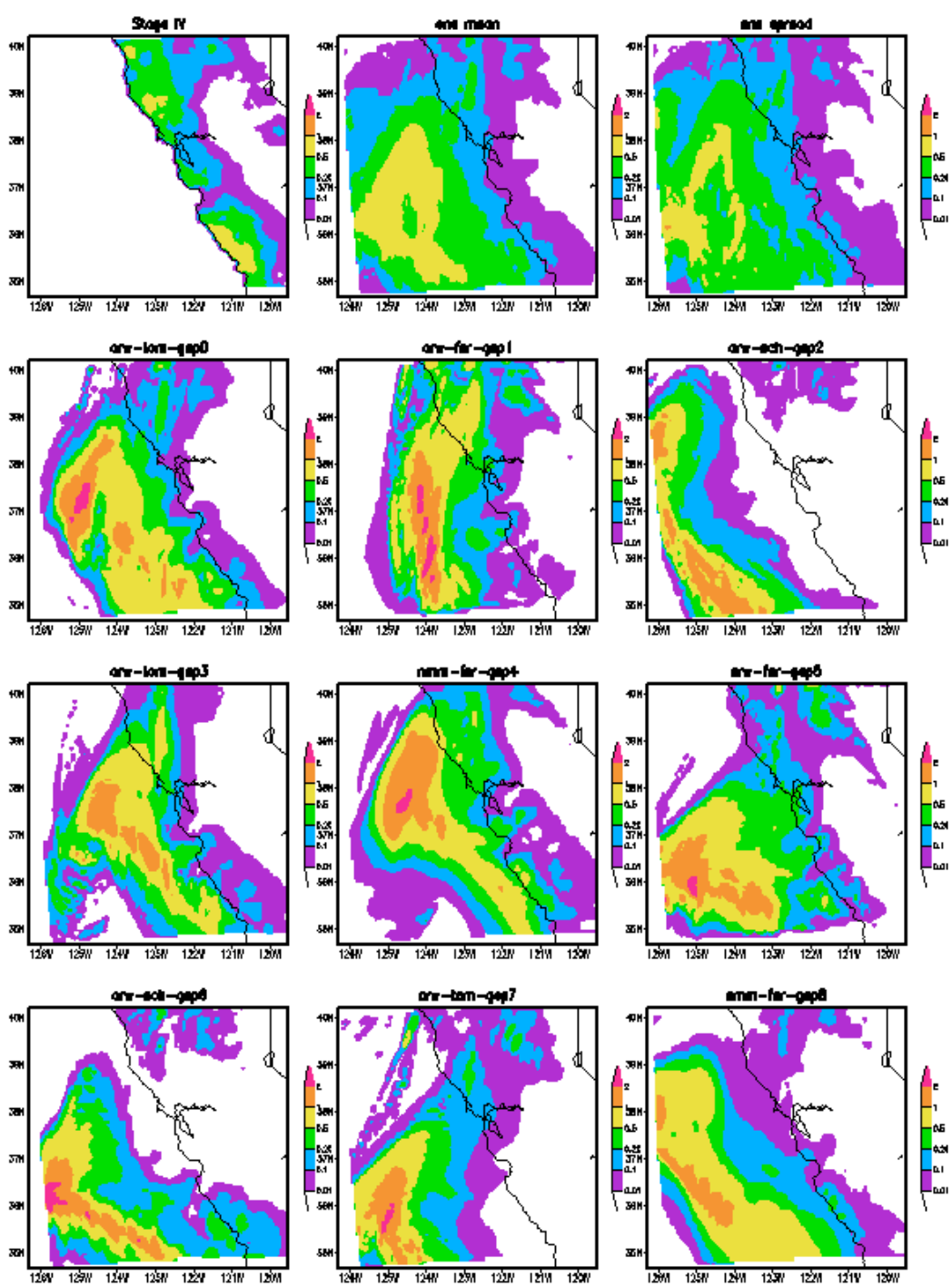
9 members:
ARW-TOM-GEP0
ARW-FER-GEP1
ARW-SCH-GEP2
ARW-TOM-GEP3
NMM-FER-GEP4
ARW-FER-GEP5
ARW-SCH-GEP6
ARW-TOM-GEP7
NMM-FER-GEP8



QPF

Example of 6-h QPF
3-km resolution

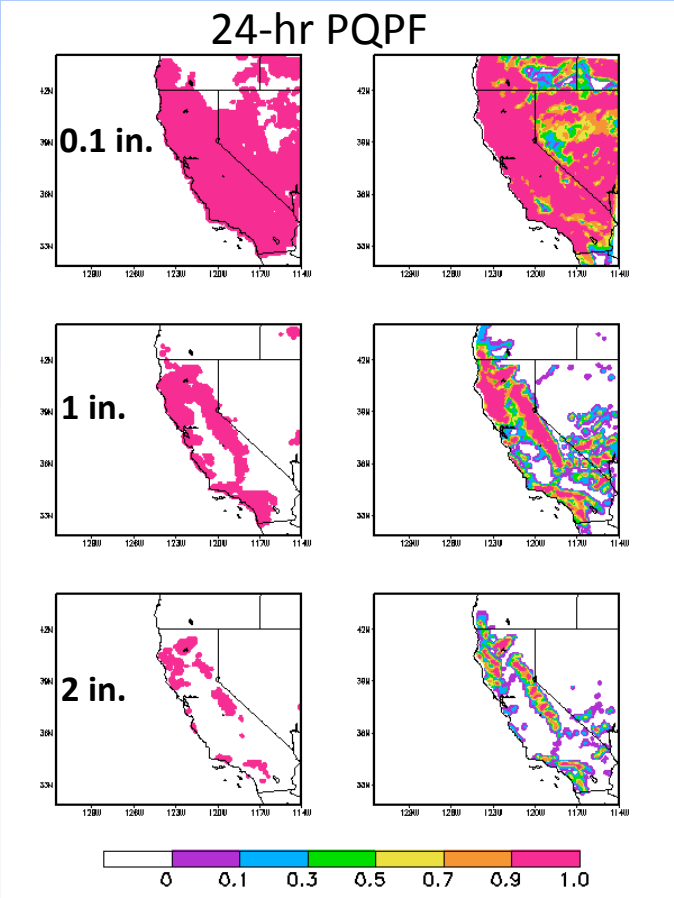
9 members:
ARW-TOM-GEP0
ARW-FER-GEP1
ARW-SCH-GEP2
ARW-TOM-GEP3
NMM-FER-GEP4
ARW-FER-GEP5
ARW-SCH-GEP6
ARW-TOM-GEP7
NMM-FER-GEP8



PQPF

Example of 24-h PQPF
9-km resolution
Probability is defined as the ratio of the number of members exceeding the precipitation thresholds of 9 total members.

Left column: observed probability based on the NCEP Stage IV data
For a given threshold, occurred event, obs probability is 1, otherwise 0.
Right column: PQPF

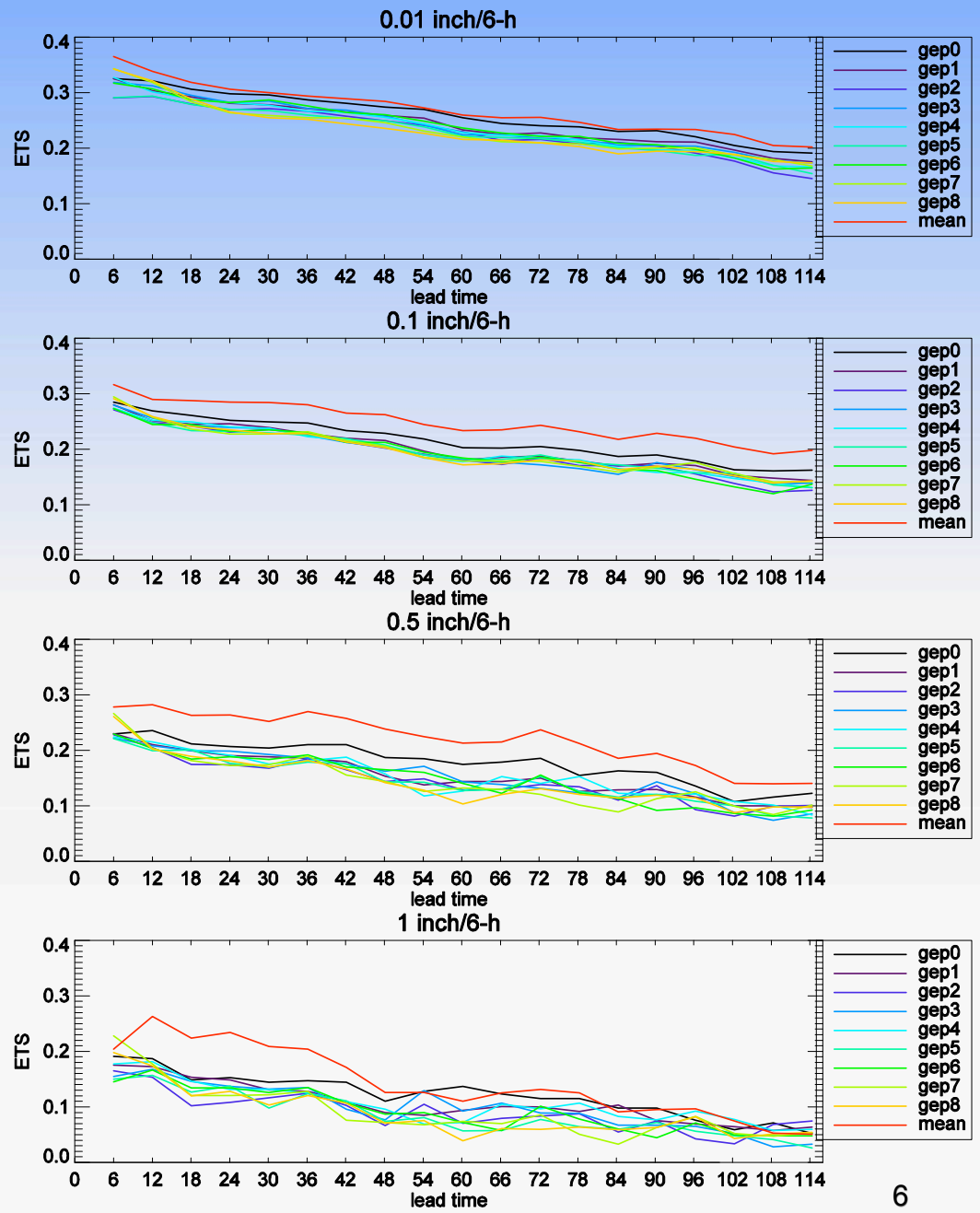


ETS of 6-h QPF

Equitable threat score (ETS) of 6-h QPF
9-km resolution
Dec 2009 - Apr 2010 (some missing data)
Verification data: Stage IV

6-h QPF verified 4 times per day (00, 06, 12, 18 UTC)
6-114 h lead times

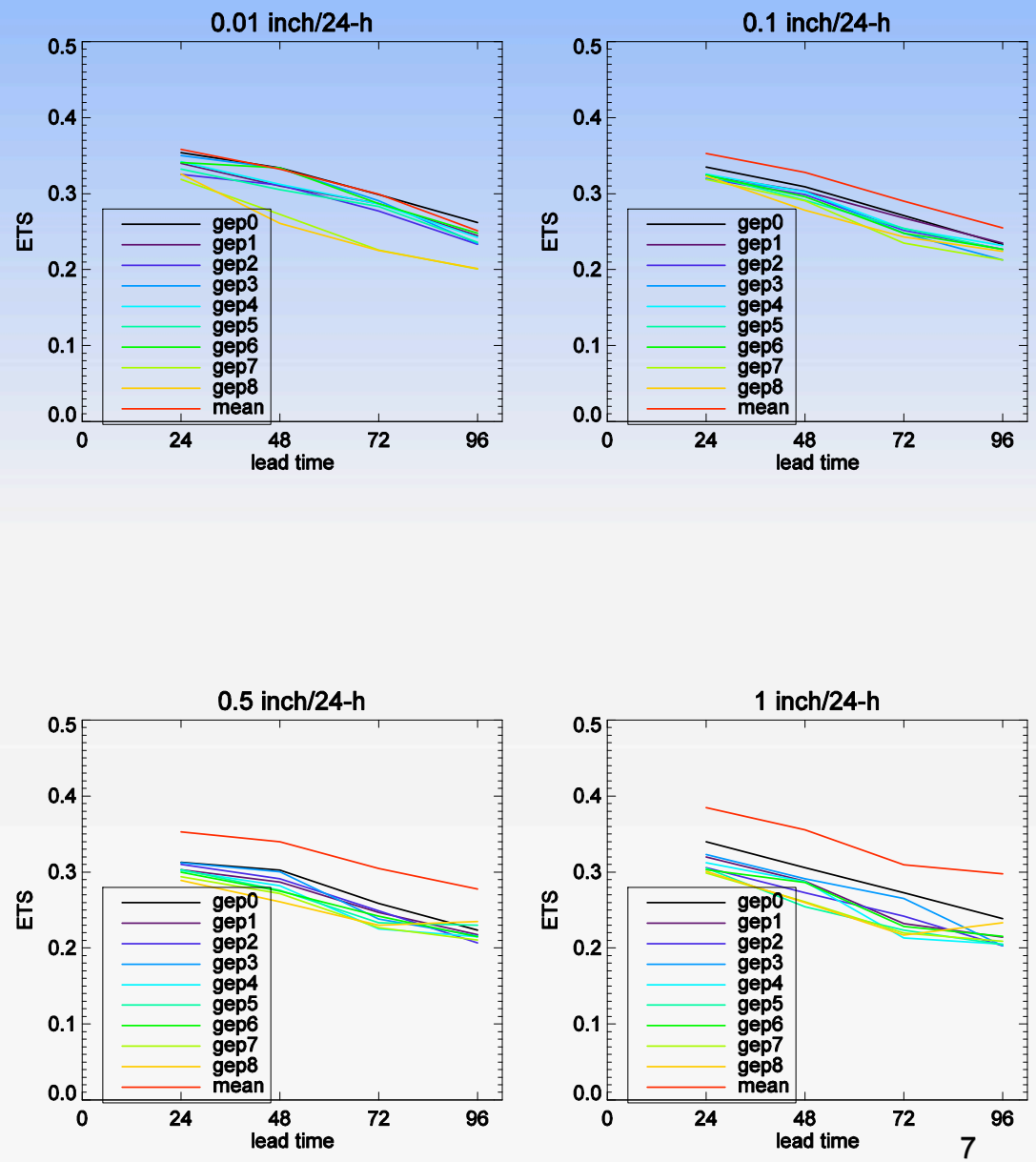
Ensemble mean is much better than individual members.
Gep0 (control) is also better.



ETS of 24-h QPF

Equitable threat score (ETS) of 24-h QPF
9-km resolution
Dec 2009 - Apr 2010 (some missing data)
Verification data: Stage IV

24-h QPF verified for 12 UTC cycle
24-96 h lead times
Ensemble mean is much better, especially for higher thresholds.

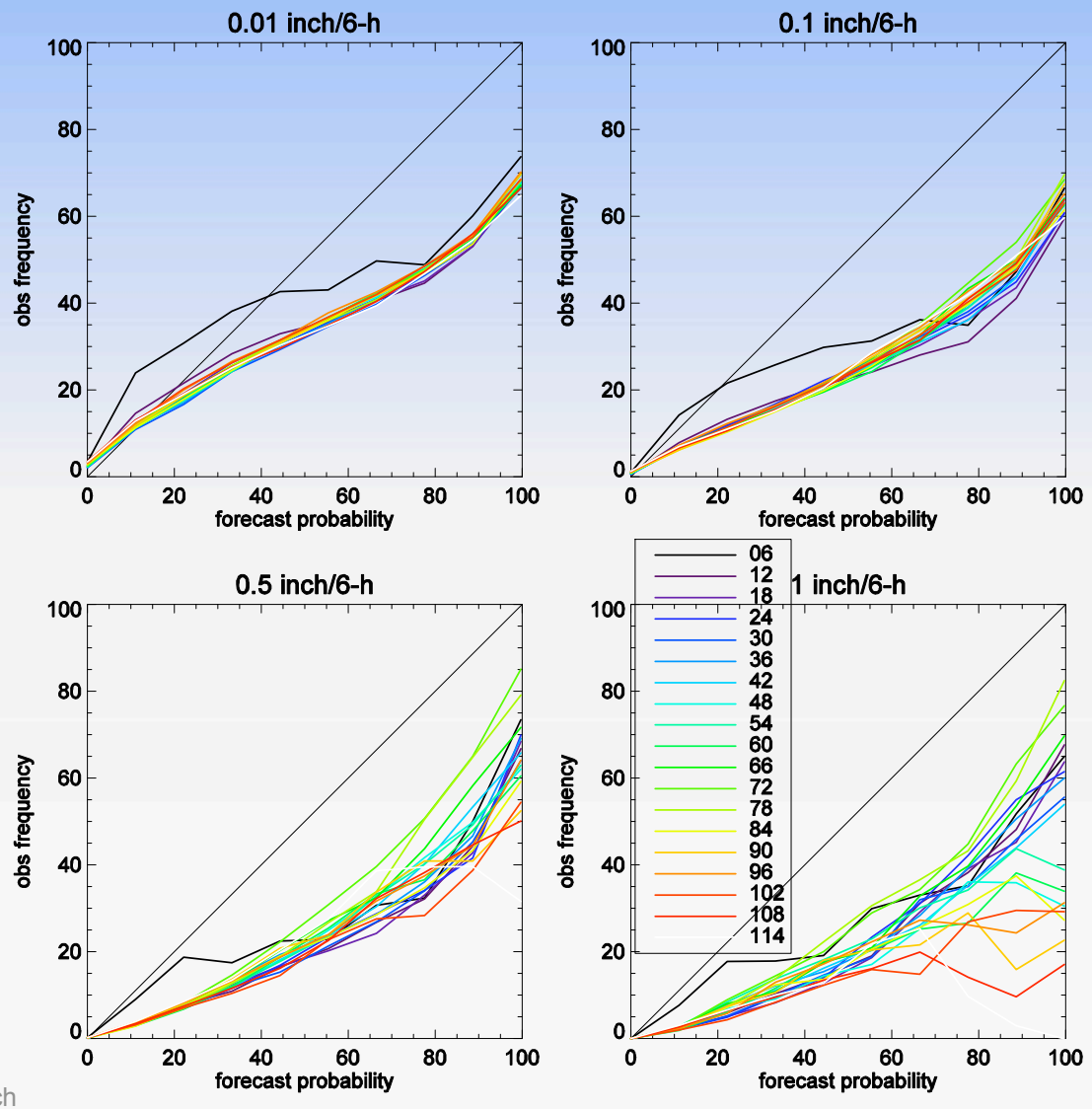


Reliability of 6-h PQPF

Reliability diagrams of 6-h PQPF
9-km resolution
Dec 2009 - Apr 2010

Observed frequency vs forecast probability
Overforecast of PQPF
Decreasing skill after 6-h lead time

Brier skill score (BSS):
Reference brier score is Stage IV sample climatology
BSS is only skilful for 6-h lead time at all thresholds and for 0.01 inch/6-h beyond 6-h lead time.

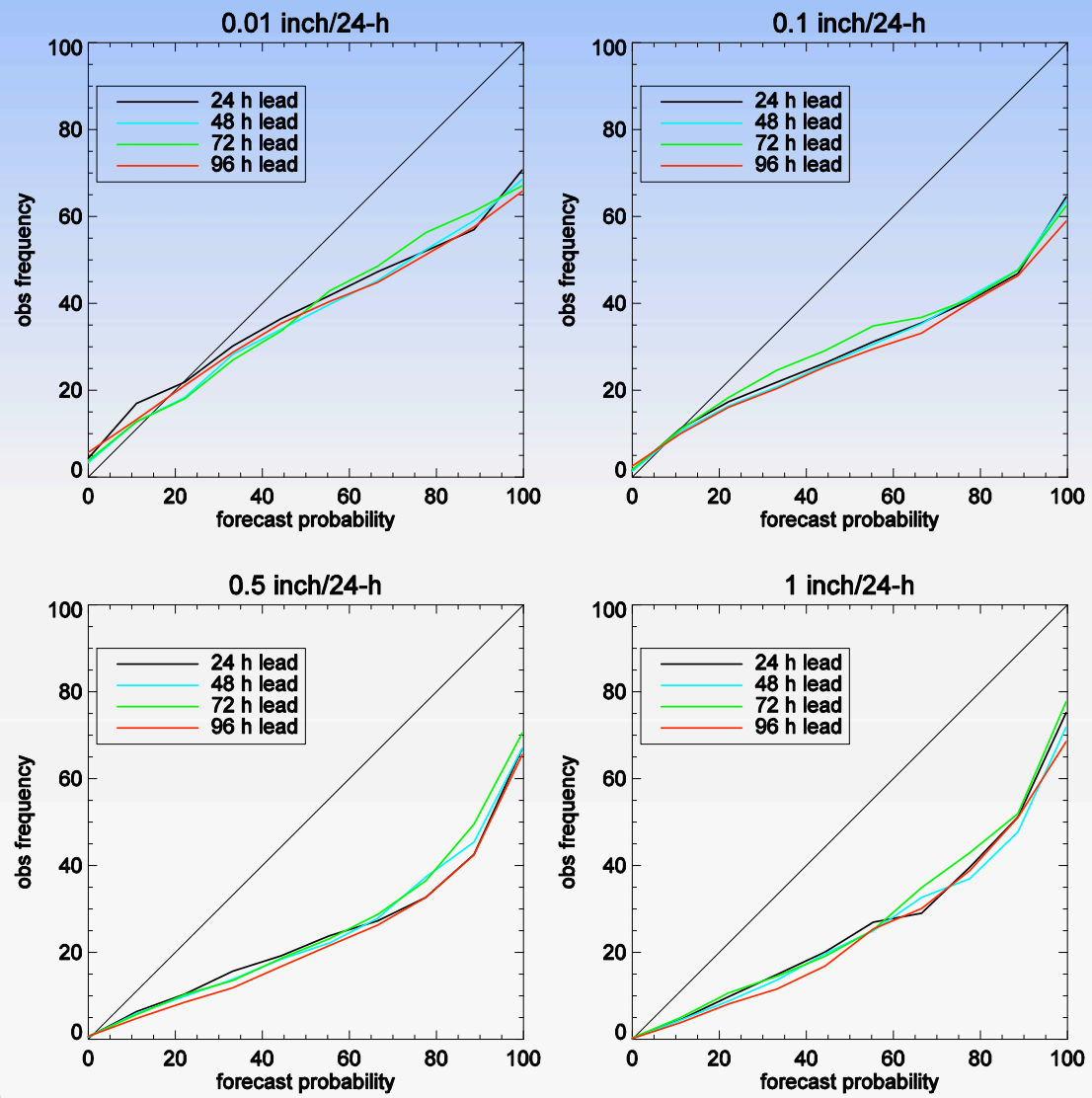


Reliability of 24-h PQPF

Reliability diagrams of 24-h PQPF
9-km resolution
Dec 2009 - Apr 2010

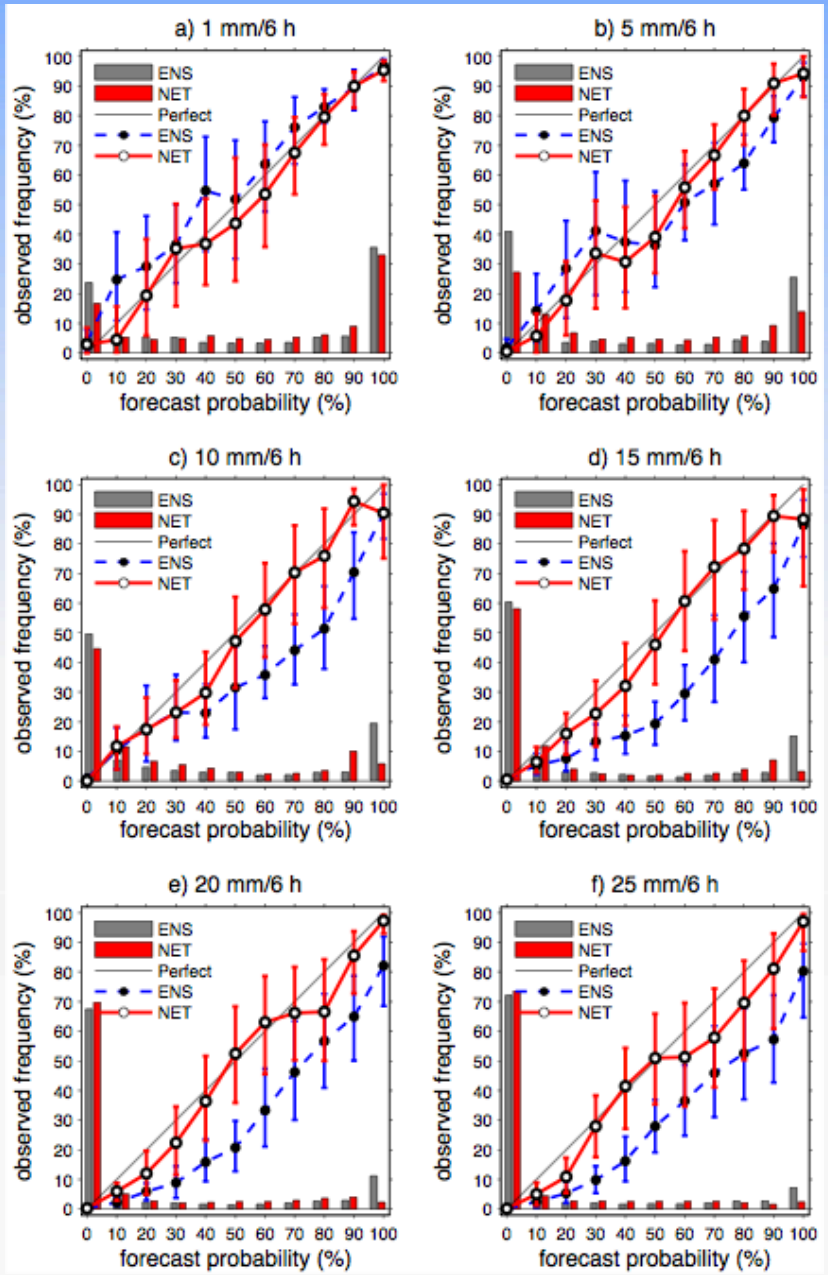
Observed frequency vs forecast probability
Overforecast of PQPF
Similar performance for different lead times

Brier skill score (BSS):
Reference brier score is Stage IV sample climatology
BSS is only skilful for 6-h lead time at all thresholds and for 0.01 inch/6-h beyond 6-h lead time.



Earlier Work – HMT Calibration of PQPF

- Calibration of probabilistic quantitative precipitation forecast (PQPF)
  - Neural network
  - Linear regression
- HMT-West multimodel time-lagged ensemble cases
- Much improved reliability



Future work

- Compare low-resolution (9 km) with high-resolution (3 km) and other models (e.g., HRRR)
- Calibration of QPF and PQPF
- Using the new precipitation analysis data (CPC + Stage IV) developed at NCEP/EMC
- Analyze other variables (e.g., temperature)
- Test for NAEFS and DTC Ensemble Testbed (DET) data
- Ensemble QPF for hydrological models (downscaling)