CPC use of the NOAA/ESRL GEFS reforecast and Improved Week-2 forecasts

Dan C. Collins

with Mike Charles, Melissa Ou, and Emily Riddle

NOAA Climate Prediction Center

Benefits of using retrospective forecasts

- Greater number of independent cases to determine forecast-observation relationships at intraseasonal timescales, compared to number of cases from recent forecast-observation pairs.
- ***Better estimation of uncertainty and reliability of probabilities***
- Bias and uncertainty estimates from recent forecast-observation pairs is always lagged in the seasonal cycle relative to real-time forecasts
  - Estimates may be associated with changing climate regime
  - Systematic errors related to lag in the seasonal cycle
  - ***Reforecast bias estimates can be centered on target dates and averaged for multiple years / regime changes***
- ***Study of model performance specific to climate modes (ENSO, MJO, AO, etc.) and regimes possible***
  - *May be essential to week 3+ forecasts*

NOAA-ESRL retrospective forecasts

- Current operational NCEP Global Ensemble Forecast System (GEFS) as of February 2012
- T254L42 (about ½ degree grid spacing) in week 1 and T190L42 (about ¾-degree) in week 2
- Daily 0Z cycle 11 ensemble members,
  - 10 perturbations + control
  - 4 cycles x 21 members per day in real-time GEFS
- 1985-2010
- Initial conditions from Climate Forecast System Reanalysis (CFSR) (2011 and real-time using GDAS)

Comparison of forecast tools from GEFS and NAEFS with and without reforecast

1. **Raw uncorrected GEFS** output
2. **Bias-corrected GEFS** using exponentially weighted moving average, adaptive bias-correction as in NAEFS from recent forecasts an observations
3. **NAEFS**: Combined bias-corrected GEFS and bias-corrected Canadian EPS using recent forecasts and observations
   - CPC adaptive precipitation CDF correction
4. **Reforecast-calibrated GEFS** using ensemble regression (Unger et al. 2009); Linear regression derived from 1984-2010 hindcast statistics

Unger, D. A., van den Dool, H., O'Lenic, E., & Collins, D. (2009). Ensemble Regression. Monthly Weather Review, 137(7).

What the Ensemble Regression corrects…

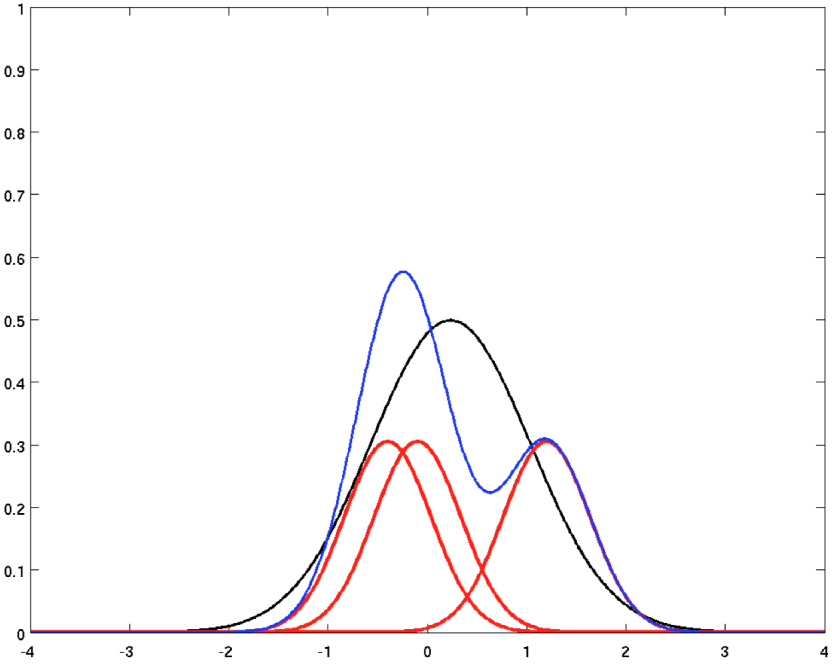
1. Removes systematic bias by removing model climatology
2. Corrects the variance of the ensemble members towards observed variance
3. *Corrects the uncertainty as indicated by the ensemble spread improving reliability of probability forecasts*

$$\sigma^2_{Error} \propto \varepsilon^2 + \sigma^2_{Spread}$$

- Accounting for model spread and skill, error variance of individual ensemble members, ε is estimated to provide width of member PDFs.

4. Damps forecasts towards observed climatology and observed climate PDF using correlation
   - Low skill forecasts resemble observed climatological probabilities

Example: Probability density functions based on linear regression:



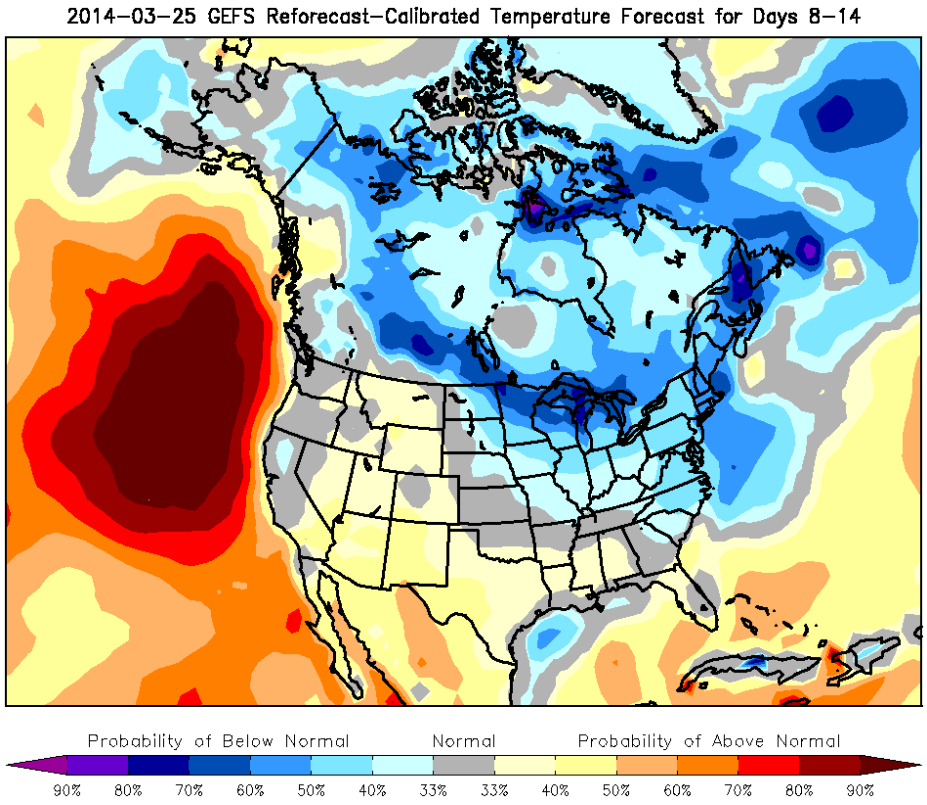
$$\sigma^2_{Error} = \sigma^2_{Obs}(1 - R^2)$$

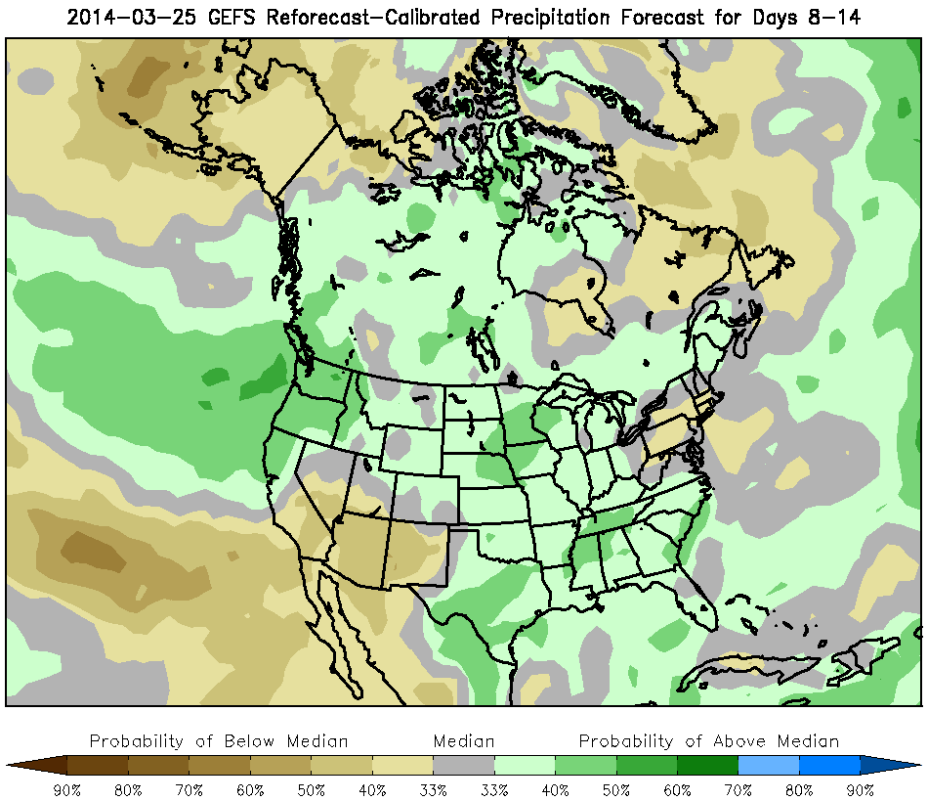
$$\varepsilon^2 = \sigma^2_{Obs}(1 - R^2_{BestMember})$$

- Individual ensemble member forecasts
- Summed ensemble regression PDF
- Regression of ensemble mean

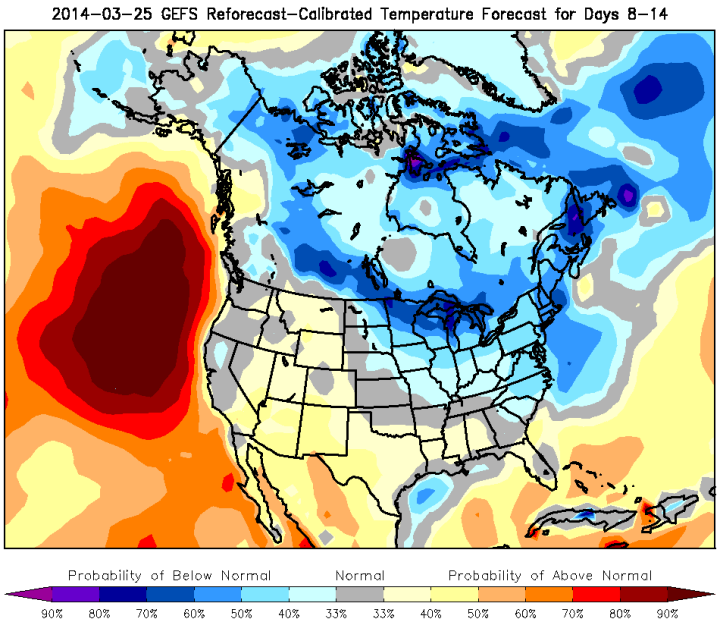
CPC week-2 temperature (left) and precipitation (right) reforecast-calibrated GEFS

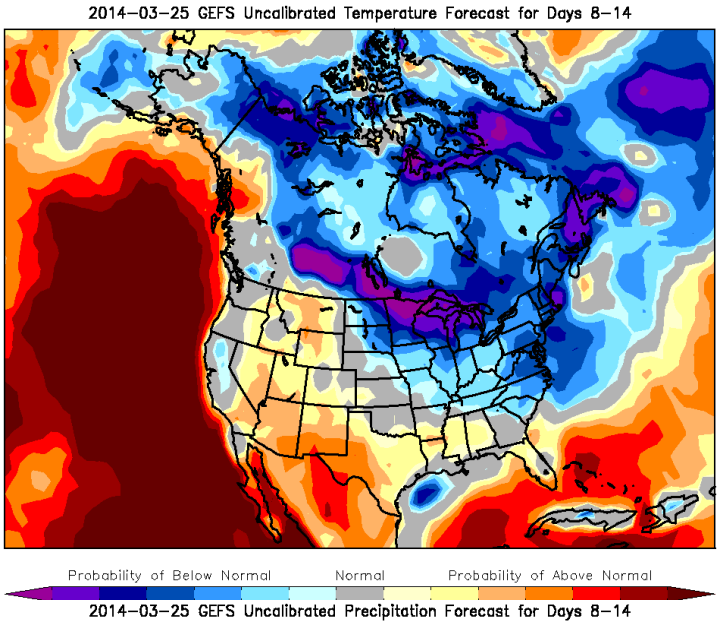
http://www.cpc.ncep.noaa.gov/products/predictions/short_range/GEFS_Reforecast/

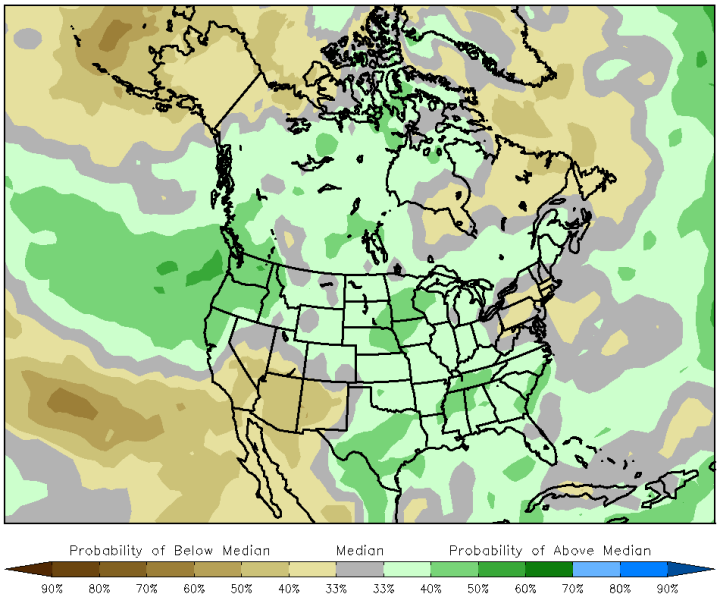


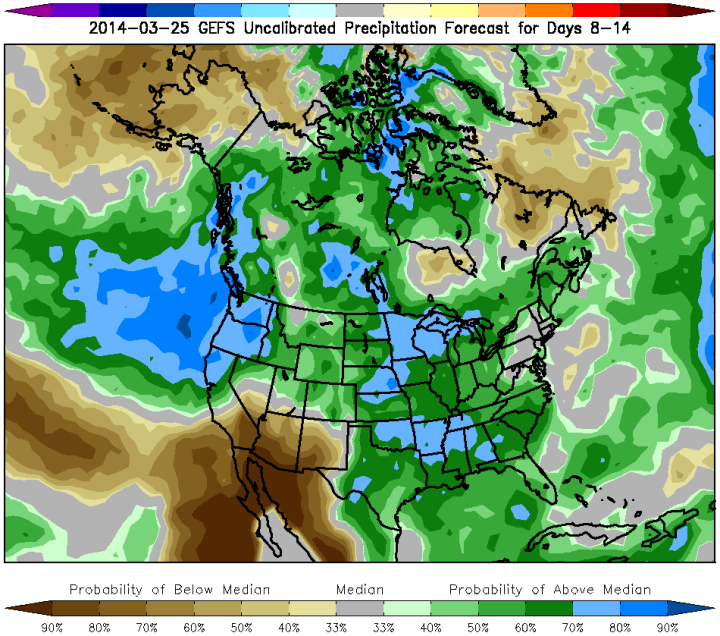


CPC week-2 temperature and precipitation reforecast-calibrated GEFS (left) and un-calibrated bias-corrected GEFS (right)

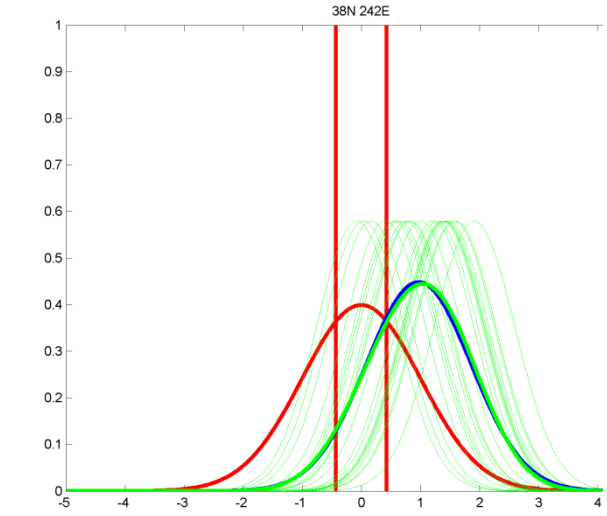


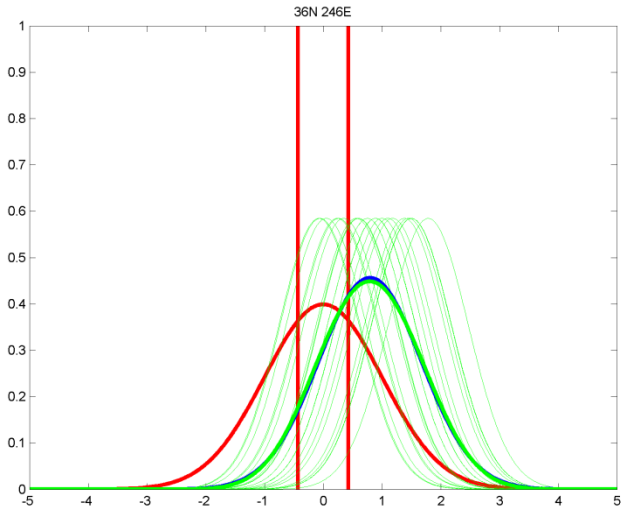


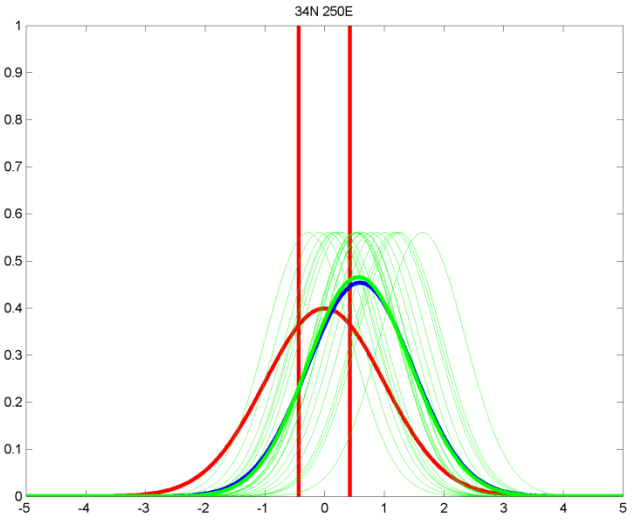


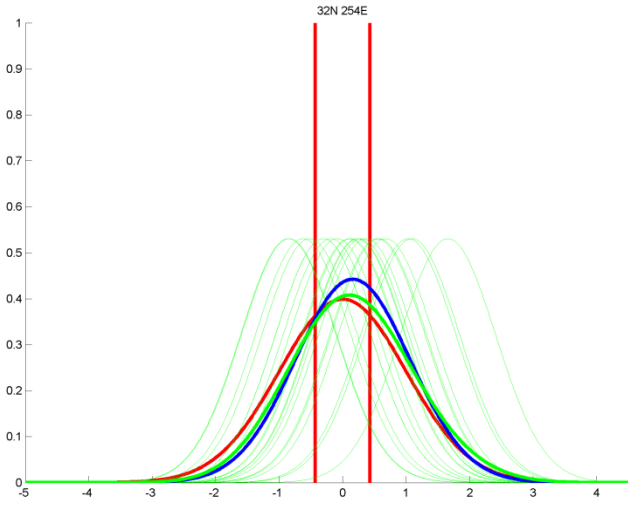


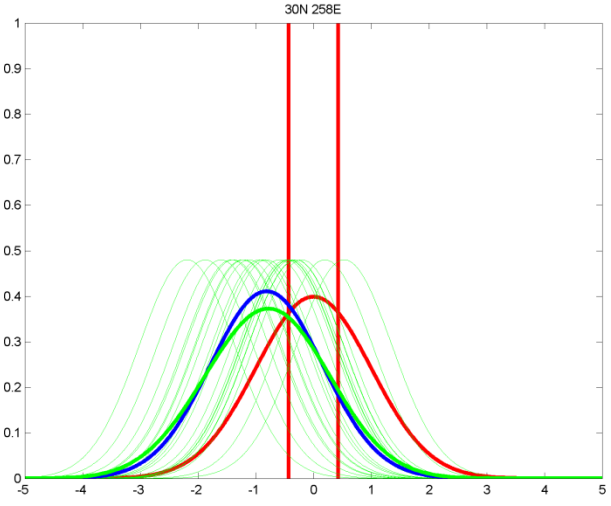
Examples from grids, of ensemble regression and regression of the ensemble mean

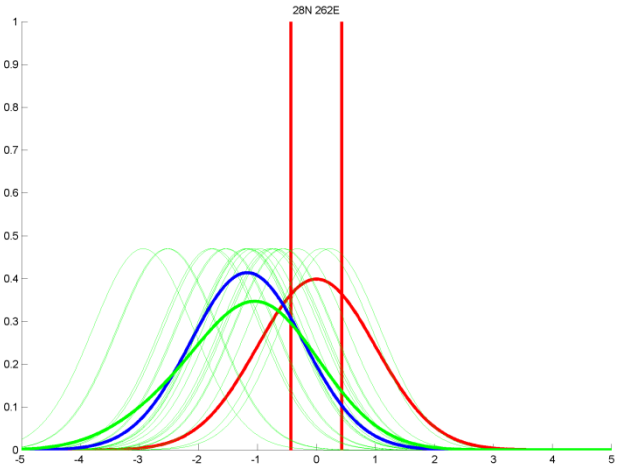






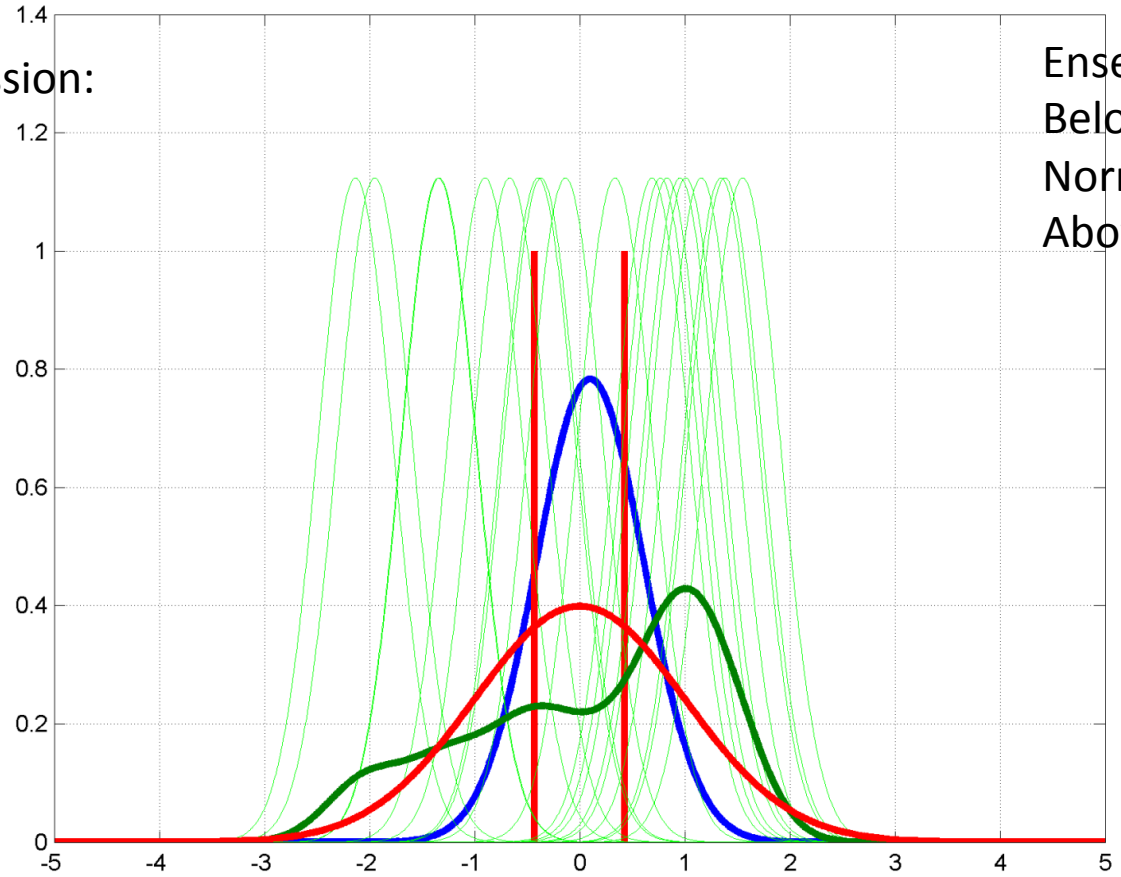






Case where ensemble regression PDF is non-normal

Mean Regression:
Below: 15%
Normal: 60%
Above: 25%

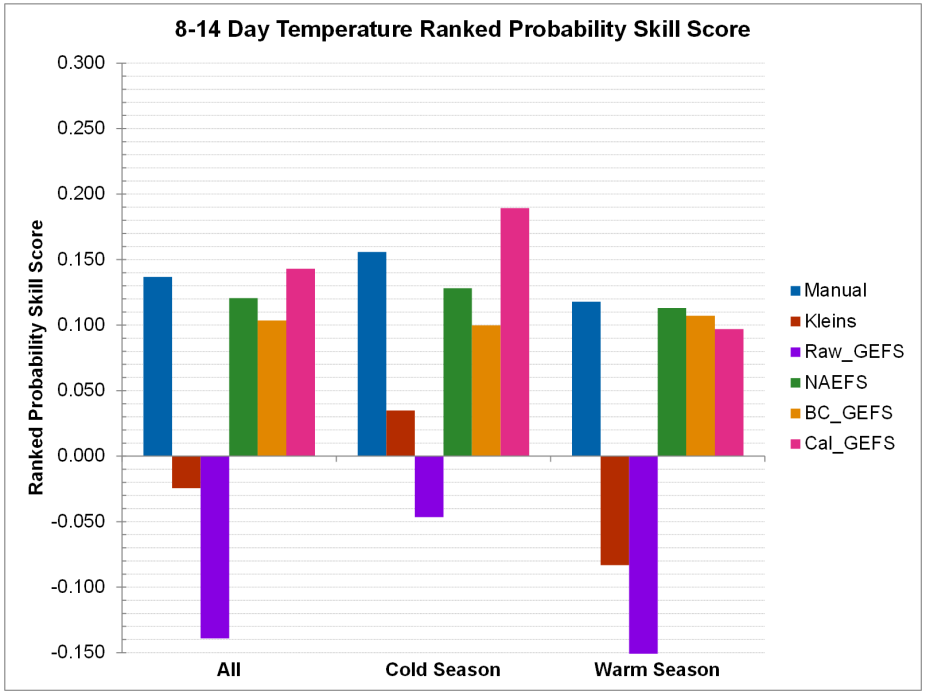


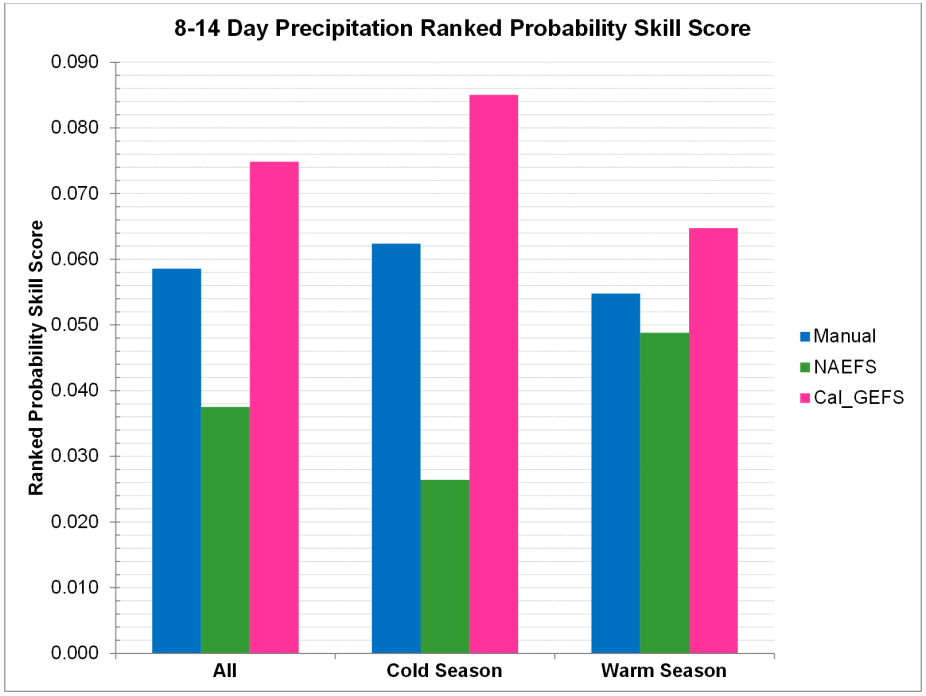
Ensemble Regression:
Below: 32%
Normal: 20%
Above: 48%

Ensemble mean anomaly: 0.3369
Regression coefficient: 0.8903
Mean forecast temp: 0.2999
Error (mean regression): 1.5516
Error of best member (ensemble regression): 1.0816
Observed standard deviation: 3.0466

Red curve: Climatological distribution
Red lines: Tercile boundaries
Blue curve: Regression of ensemble mean pdf
Dark green curve: Ensemble regression pdf
Light green curves: pdfs of individual members

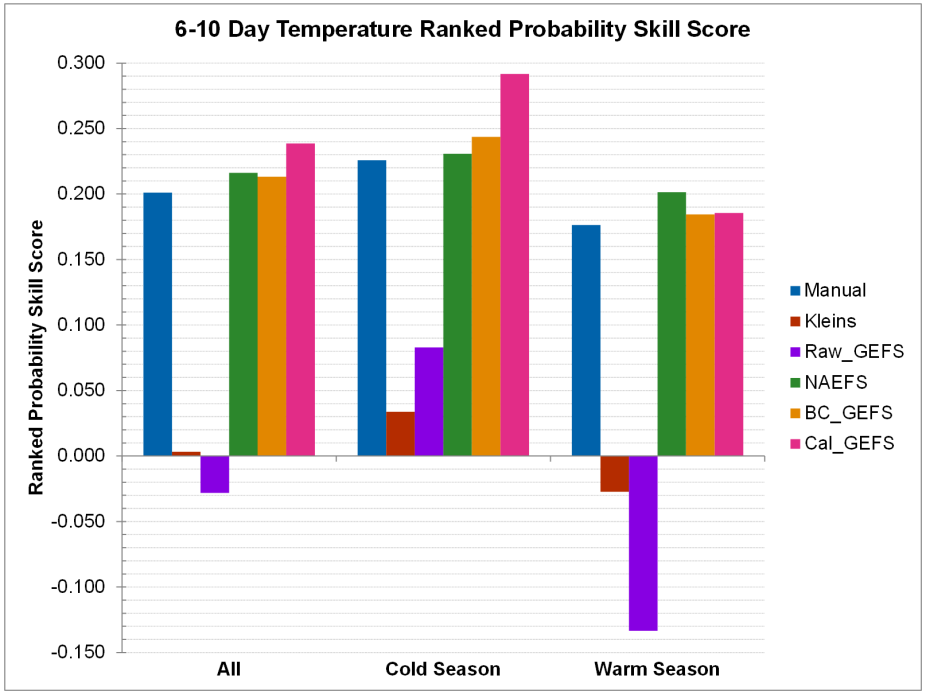
Week-2 Temperature (left) and Precipitation (right) Rank Probability Skill Scores (RPSS) of reforecast-calibrated GEFS compared to NAEFS and GEFS

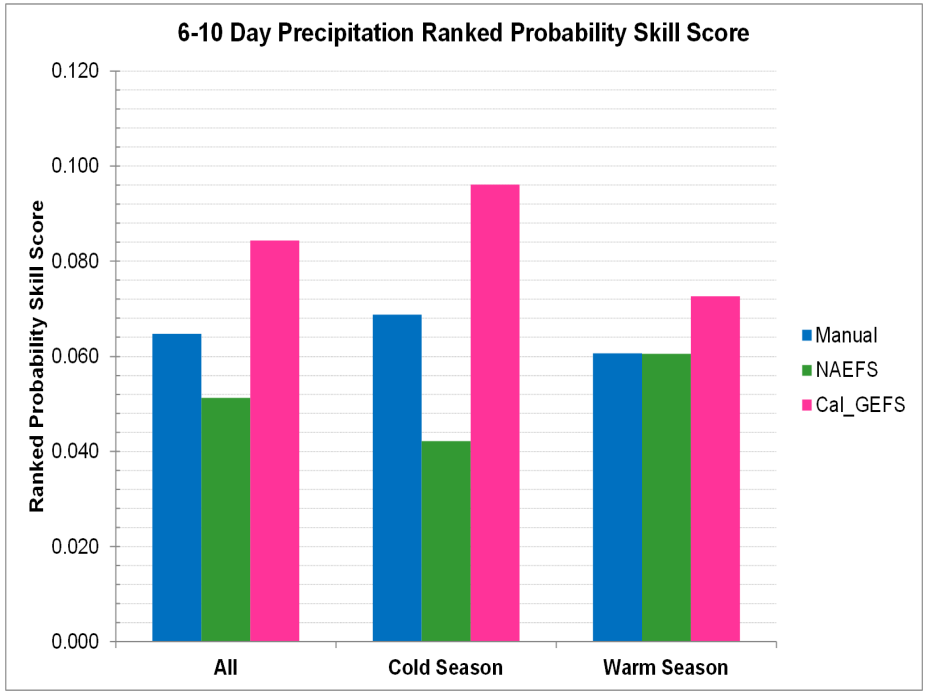




Comparison of tools for CPC week-2 forecasts including official manual forecasts, raw GEFS, bias-corrected GEFS component of NAEFS, NAEFS MME, and reforecast-calibrated GEFS.

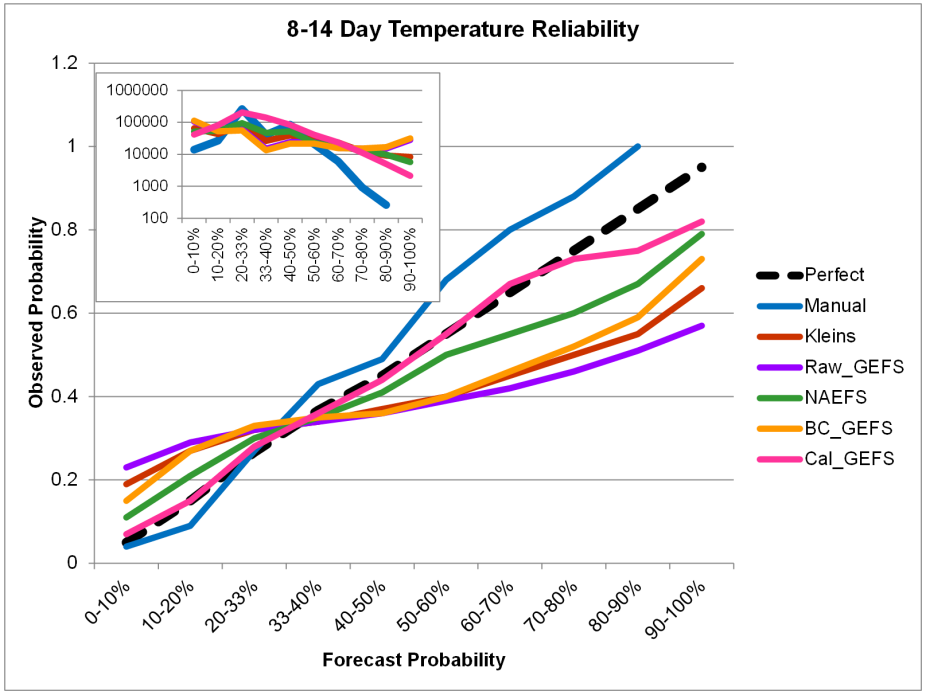
6-10 day Temperature (left) and Precipitation (right) Rank Probability Skill Scores (RPSS) of reforecast-calibrated GEFS compared to NAEFS and GEFS

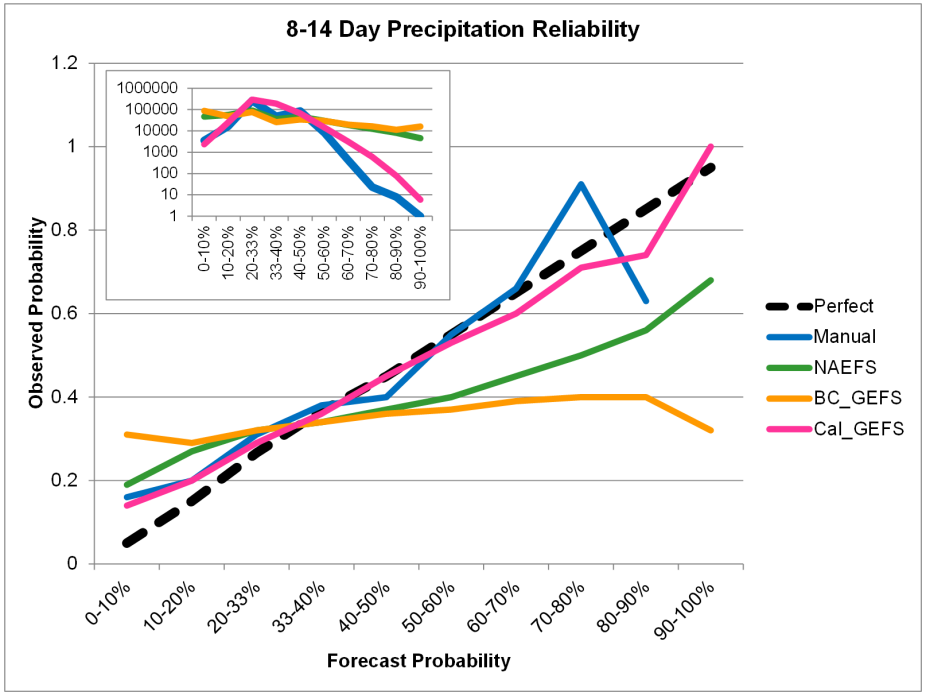




Comparison of tools for CPC week-2 forecasts including official manual forecasts, raw GEFS, bias-corrected GEFS component of NAEFS, NAEFS MME, and reforecast-calibrated GEFS.

Week-2 reliability of reforecast-calibrated GEFS probabilities compared to NAEFS, bias-corrected GEFS and uncorrected GEFS (3 years 2011-2013)





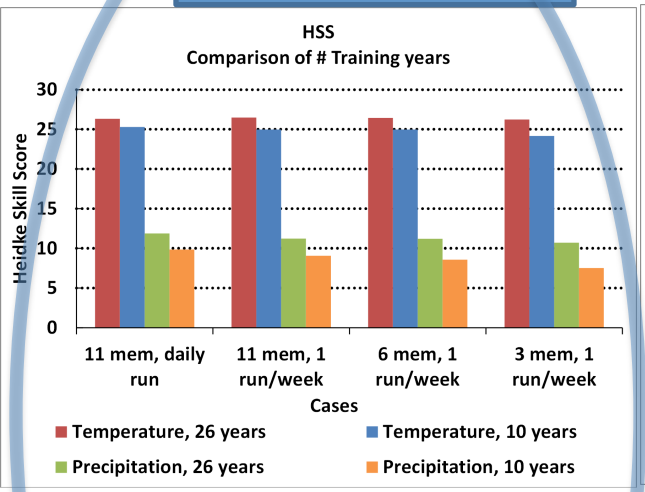
Comparison of reliability of CPC week-2 probability forecasts including official manual forecasts, raw GEFS, bias-corrected GEFS component of NAEFS, NAEFS MME, and reforecast-calibrated GEFS.

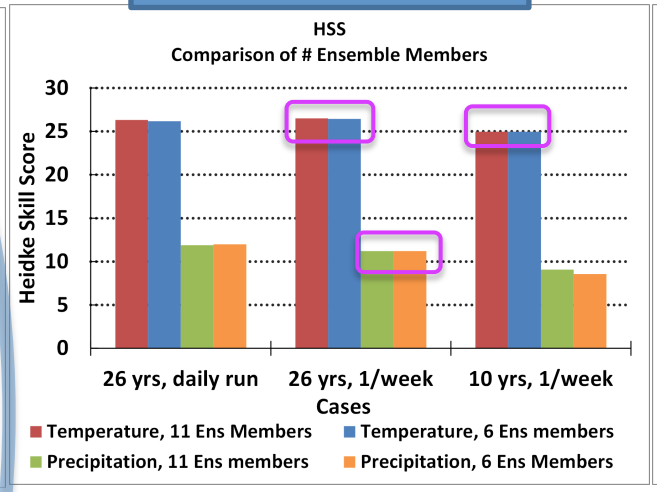
Sampling a smaller reforecast

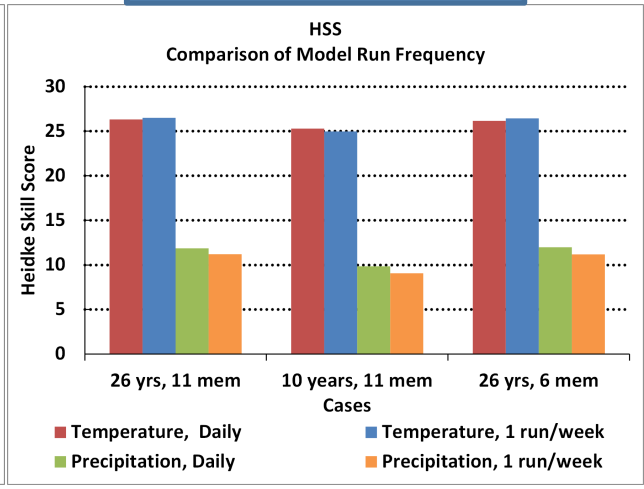
26 vs. 10 Training Years

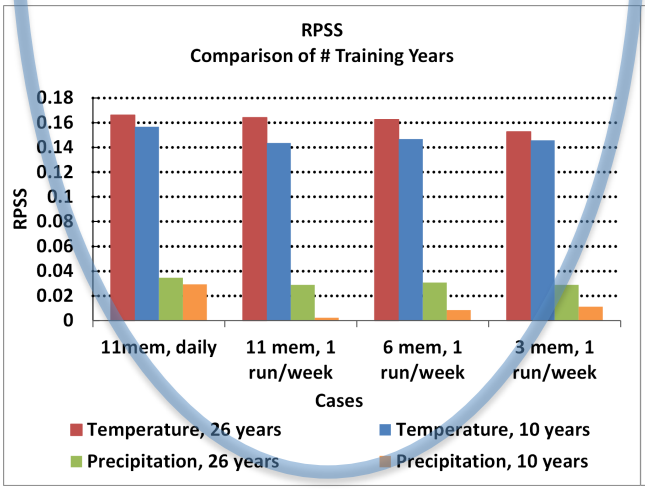
11 vs. 6 Members

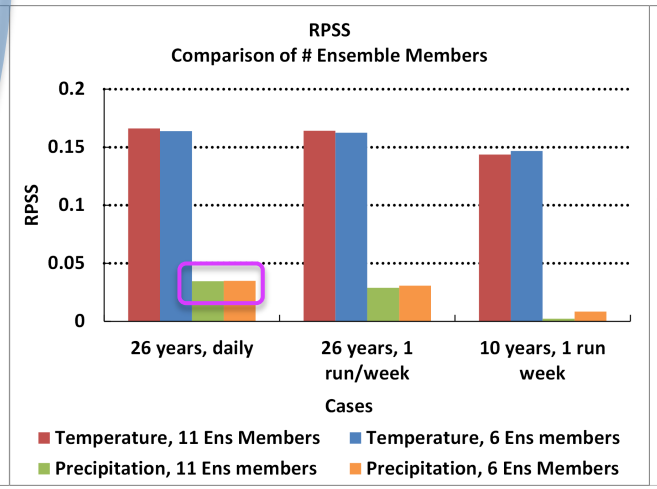
Daily vs. 1 run

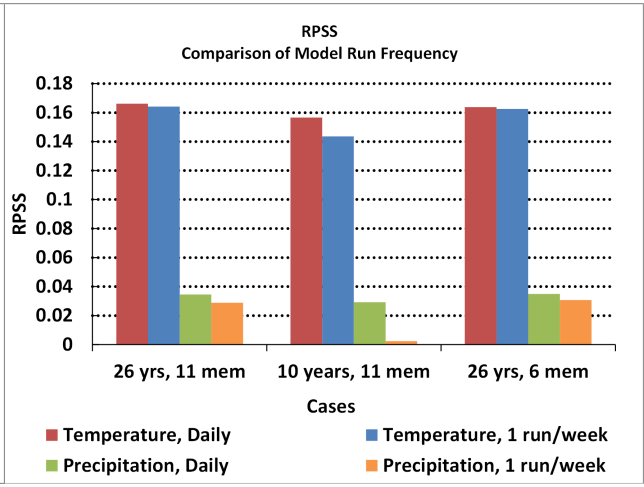






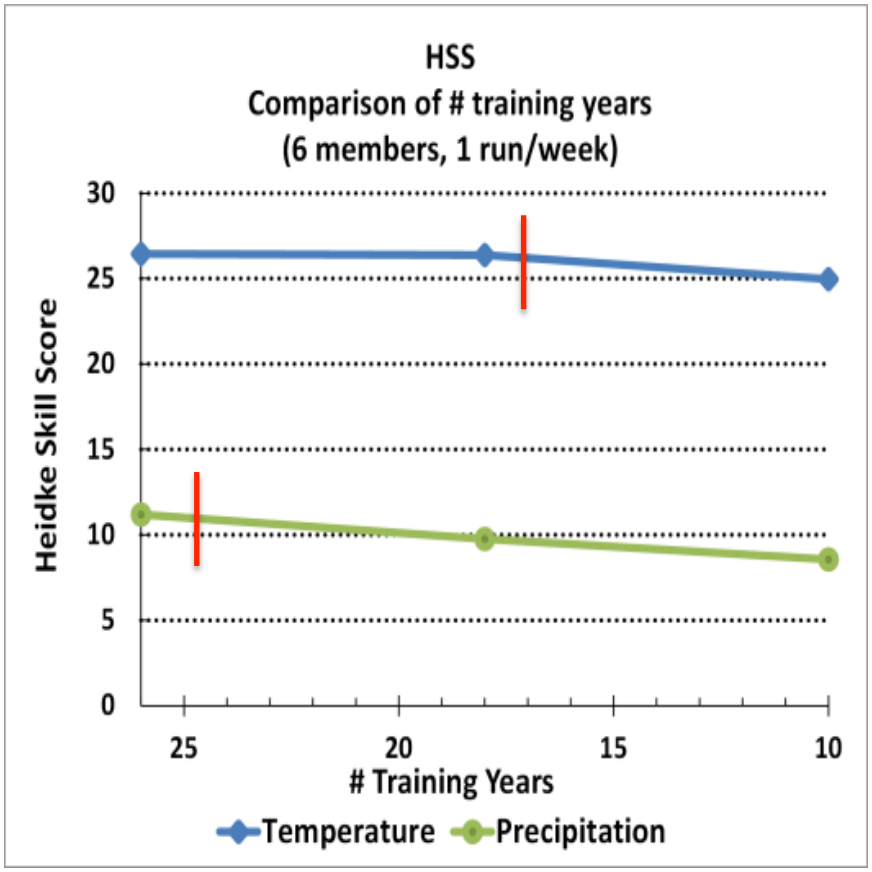


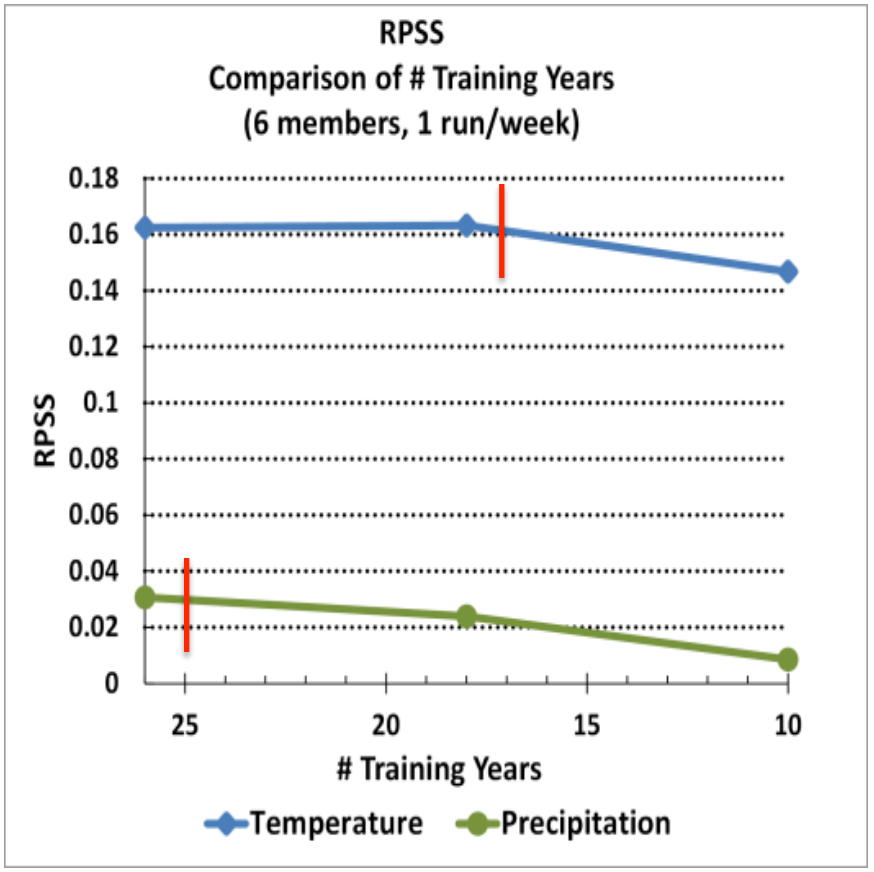




Results of reforecast sample size tests for week-2 temperature and precipitation show reducing the number of years in the reforecast produces the largest drop in skill. Reducing the ensemble size from 11 to 6 members had little or no impact.

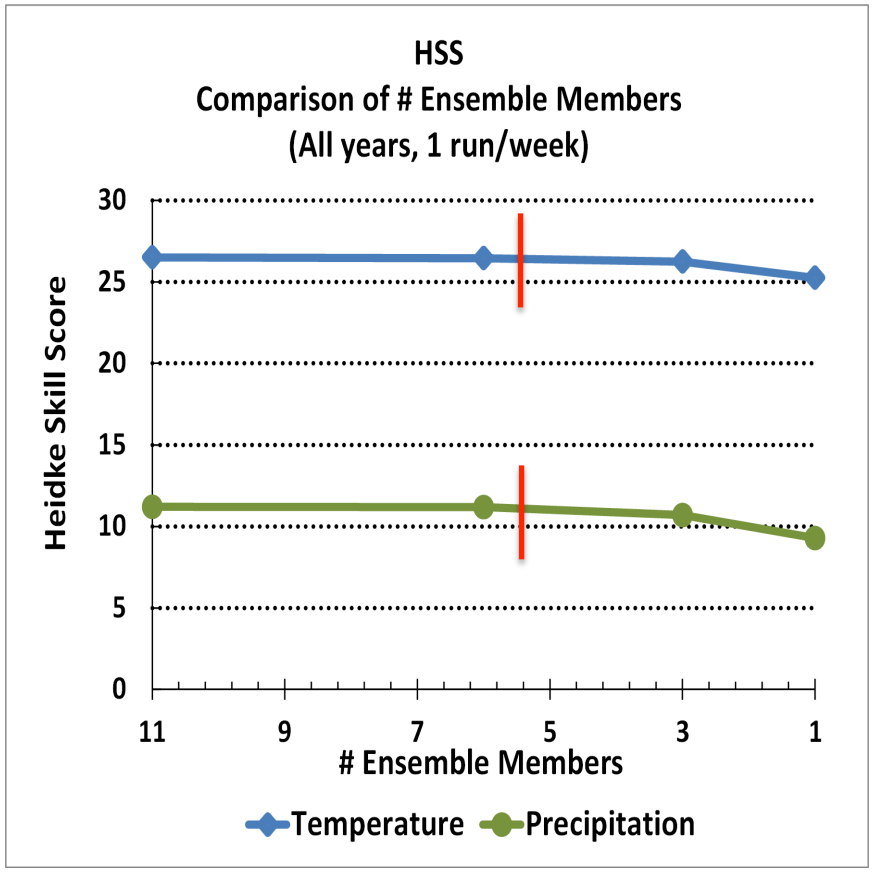
Testing the impact on week-2 temperature and precipitation skill of a reduced number of **years** in reforecast sample

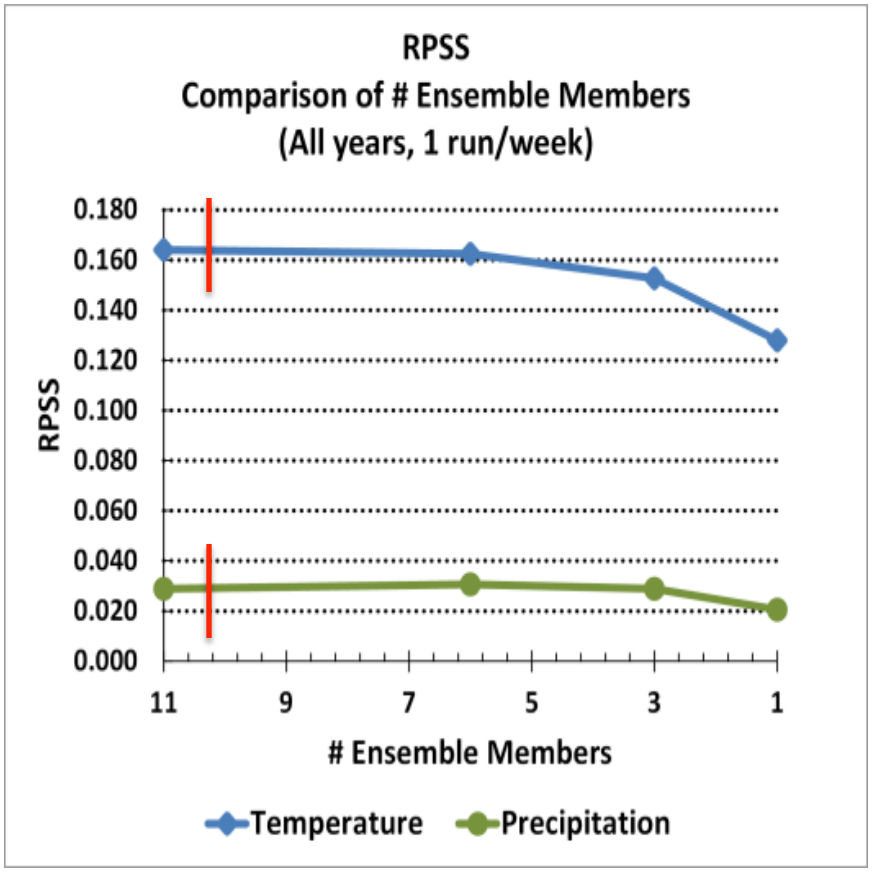




Using 26, 18 and 10 years of reforecasts (1984-2010) to generate statistics for calibration of 2011-2013 GEFS forecasts. Heidke Skill Score (left) and RPSS (right) show ***loss of skill for precipitation forecasts with reduction from full 26 years.***
Red lines mark first reduced sample with significantly lower skill.

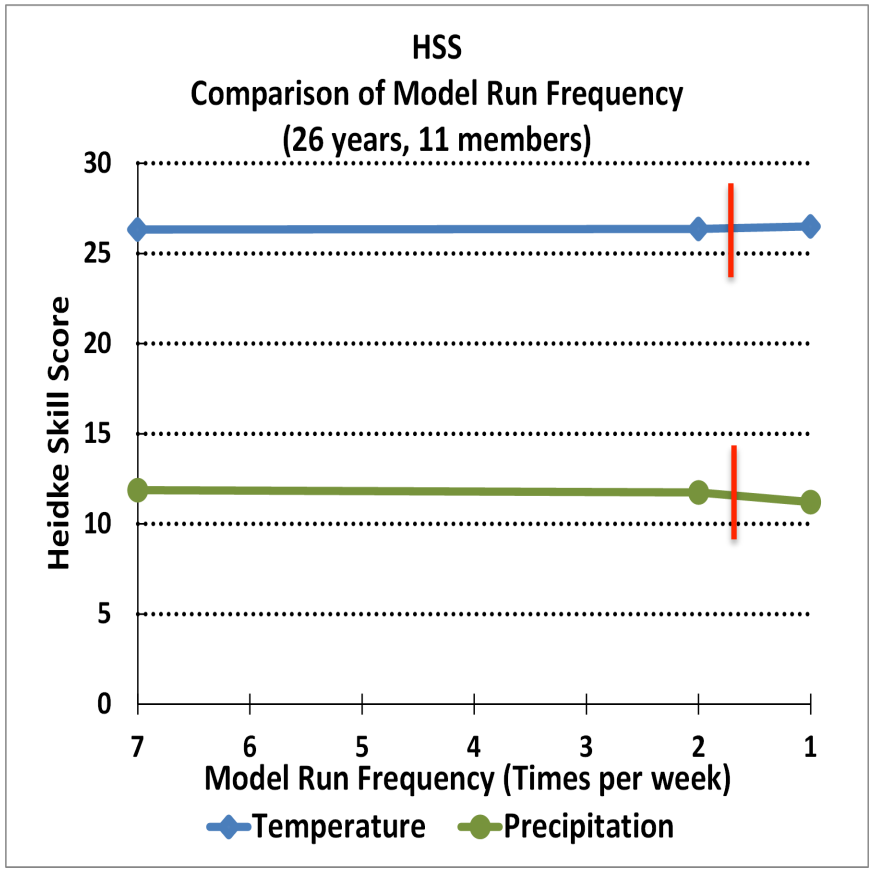
Testing the impact on week-2 temperature and precipitation skill of a reduced number of **ensemble members** in reforecast sample

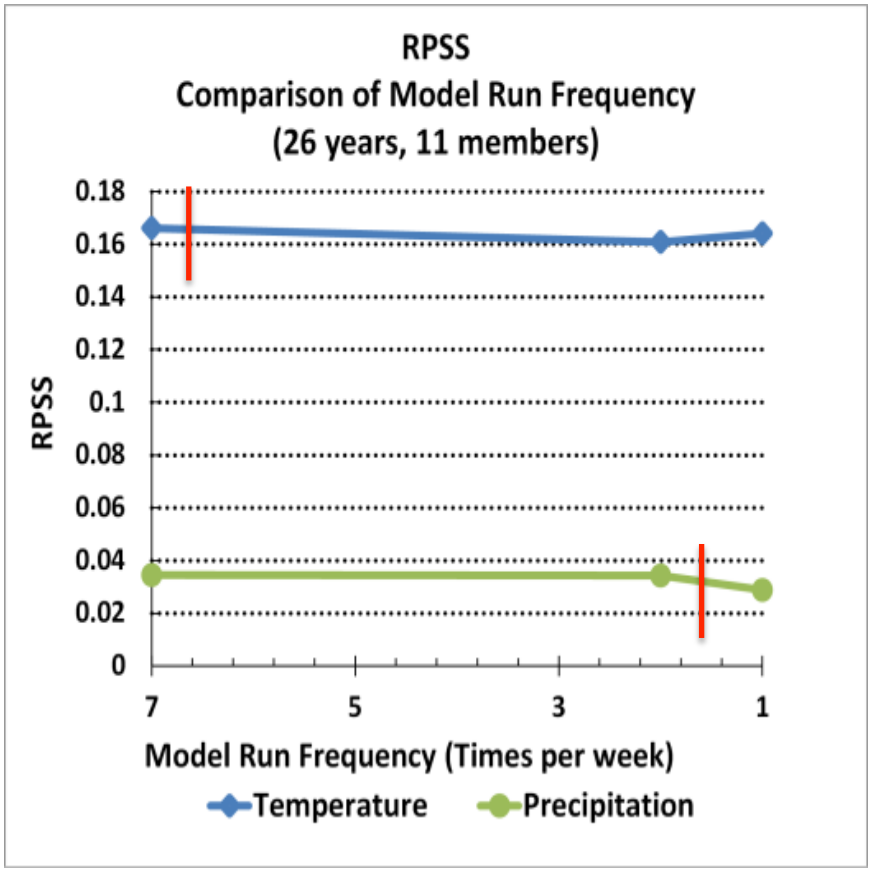




Using 11, 6, 3 and 1 ensemble members of reforecasts (1984-2010) to generate statistics for calibration of 2011-2013 GEFS forecasts. Heidke Skill Score (left) and RPSS (right) show ***little or no loss in skill with reduction from 11 to 6 ensemble members.***

Testing the impact on week-2 temperature and precipitation skill of a reduced frequency of runs from daily to weekly

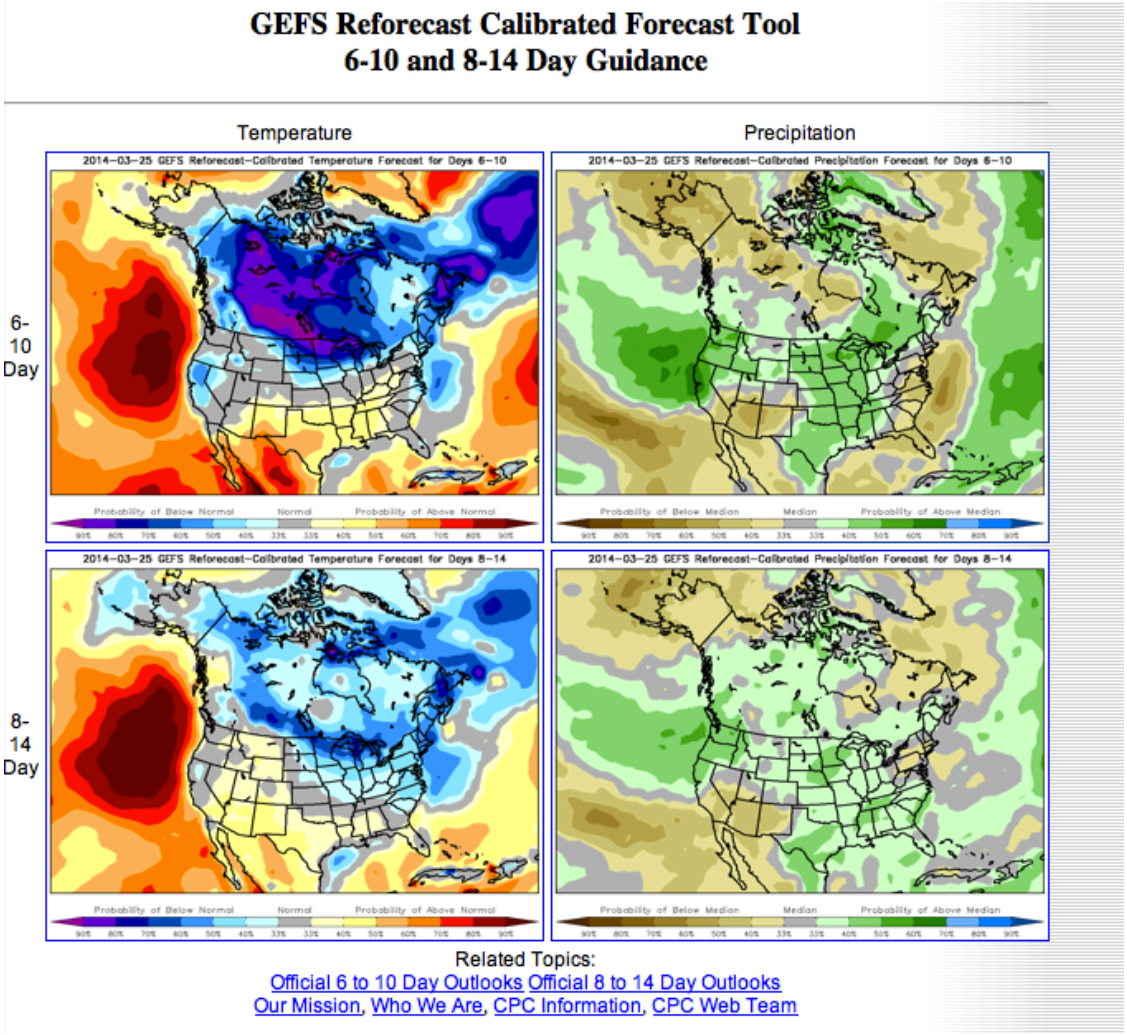




Using 7, 2 and 1 run per week (1984-2010) of reforecasts to generate statistics for calibration of 2011-2013 GEFS forecasts. Heidke Skill Score (left) and RPSS (right) show ***little or no loss in skill with reduction from daily to twice-weekly runs.***

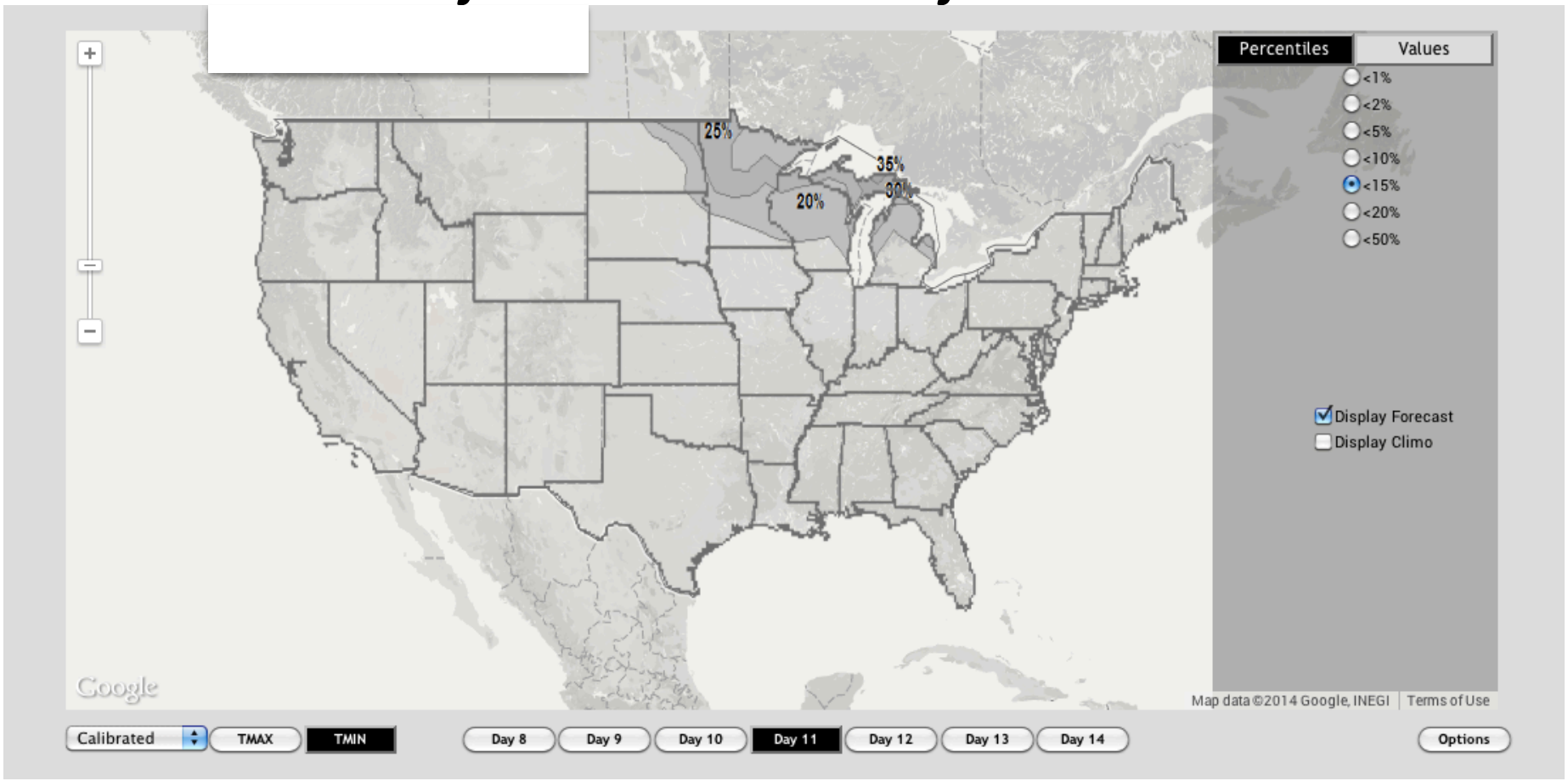
CPC 6-10 day and week-2 temperature (left) and precipitation (right) reforecast-calibrated GEFS webpage

http://www.cpc.ncep.noaa.gov/products/predictions/short_range/GEFS_Reforecast/



CPC week-2 daily temperature extremes

Probabilistic temperature extremes forecast tool requires reforecast to corrected full PDF



Probability of day 11 daily minimum temperatures below the lower 15th percentile from climatology.

Summary and Conclusions

- Reforecast-calibrated GEFS for week-2 temperature and precipitation is the most skillful CPC forecast tool overall.

- Estimation of the relationship between the model spread and uncertainty is essential to reliability of CPC's probabilistic forecasts.

- Forecast skill of reforecast-calibrated GEFS is most sensitive to a reduction in the number of available years in the reforecast system. *The more years the better.*

- *CPC probabilistic daily extremes forecasts in the extended range require a reforecast system.*

- *Use of the GEFS reforecast system will be a key component of CPC extended range multi-model consolidation, along with (hopefully) use of the ECMWF and Environment Canada reforecast systems.*