Use of ensemble forecasts to assess the predictability of subseasonal forecasts

Dan C. Collins

Acknowledgements: Emily E. Riddle , Mike Charles, Melissa Ou and David Unger

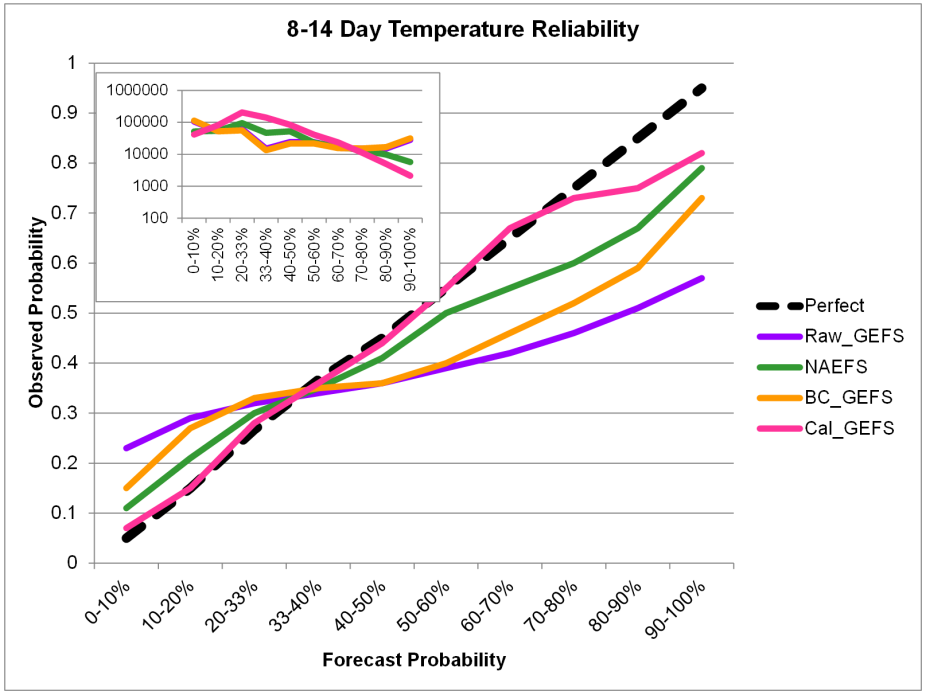
NOAA Climate Prediction Center (CPC)

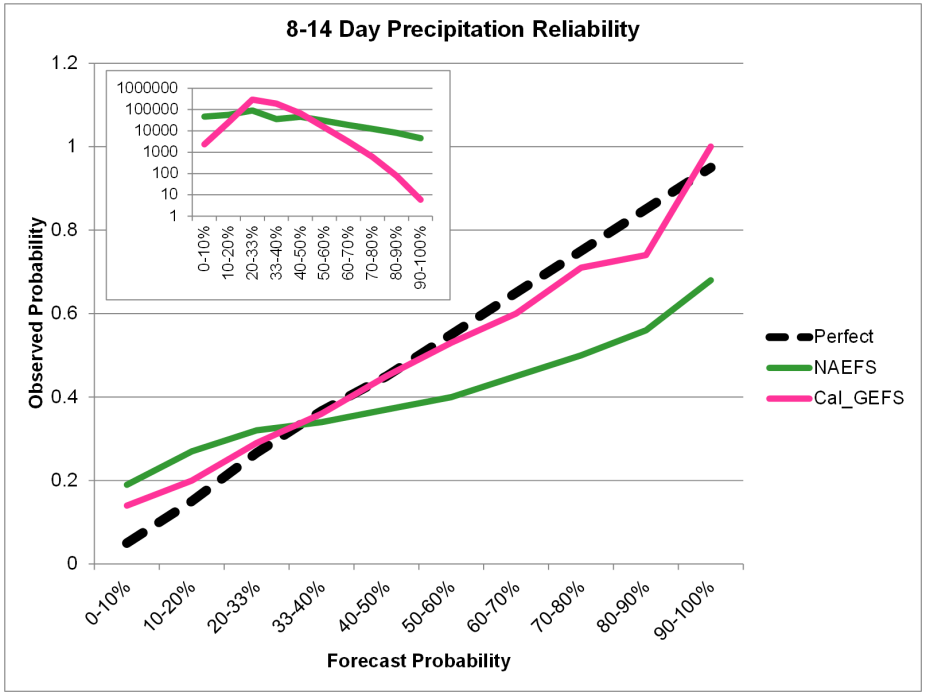
Using NOAA-ESRL retrospective forecasts:

Retrospective forecasts for week-2, 2-meter temperature from the Global Ensemble System for 1985-2010 used to estimate model climatology, including the mean and variance, correlation with observations.

- 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)

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.

Some ideas up front

- The predictable signal can be estimated from the variance of the ensemble mean (magnitude of anomalies) and ensemble spread is an estimate the uncertainty.

  (*There are various definitions of predictability.*)

- Because variance is additive, signal variance plus spread should approximately equal variance in observations; This means our estimates of predictability are mean and variance bias adjusted.

  (*Should both signal and noise variance be adjusted? Or only noise variance?*)

  - Predictability is determined from the model-predicted correlations calculated from the signal and ensemble spread.

- Comparing this to the realized correlation of forecasts to observations tells us if the current prediction capacity matches predictability estimates.
- Trends can be fit to the predictability, both model-estimated and realized, as well as to the difference between the model climatology and the observed climatology, i.e. bias.
- Are trends in bias an indication that systematic bias between the ensemble mean and observations is changing?
  - Some of the "bias" changes can be seasonal climate-state dependent!

  (*Both the mean climate state and the observational analysis change*.)

Using Ensembles for Predictability Estimates

What do we mean by predictability?

- The **ensemble mean** is an estimate of the predictable component of the future state, also known as *signal* ***S***

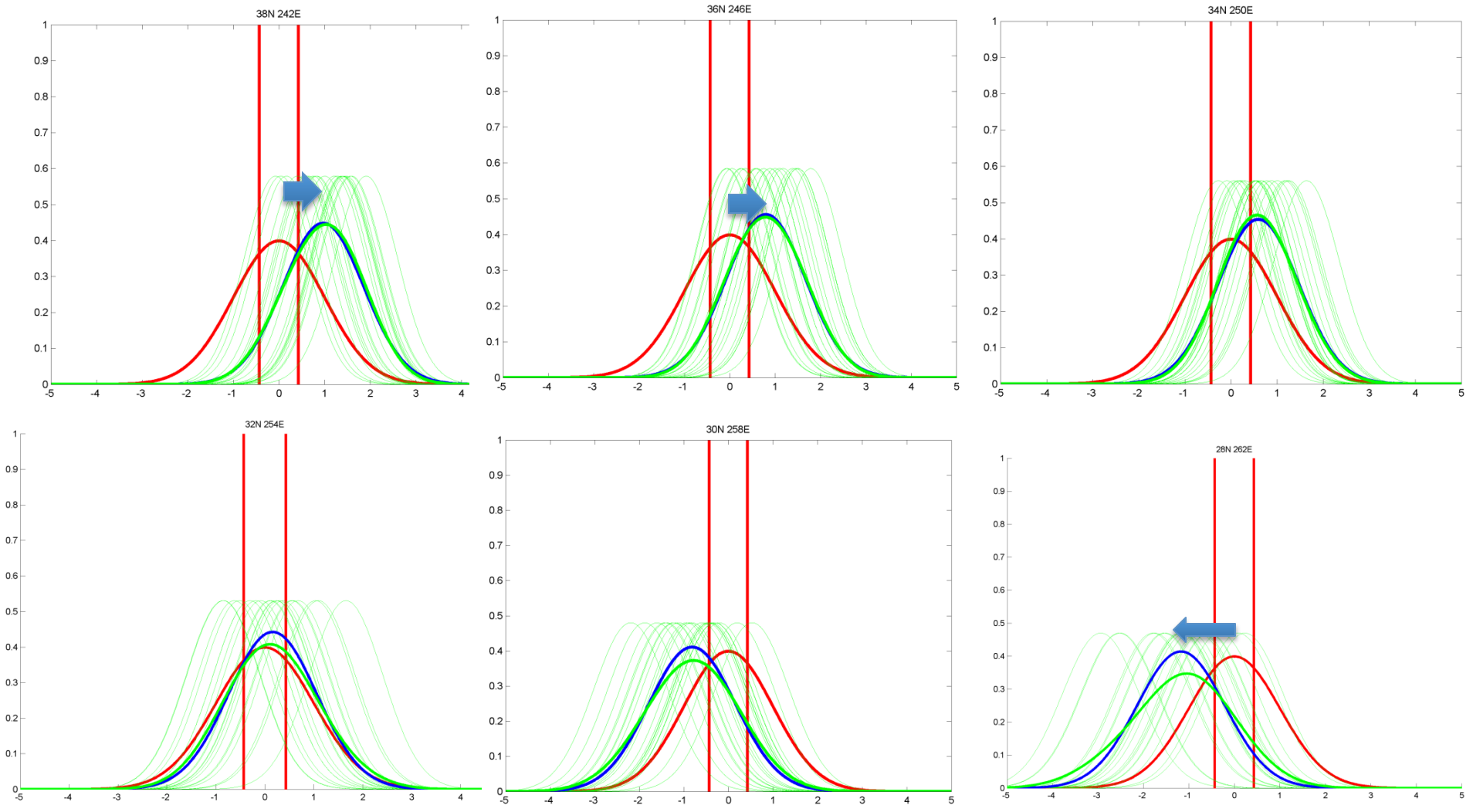
$$x_t = \langle x \rangle_t + x'_t$$

- The **ensemble spread** about the mean is an estimate of the unpredictable *noise*, **N**

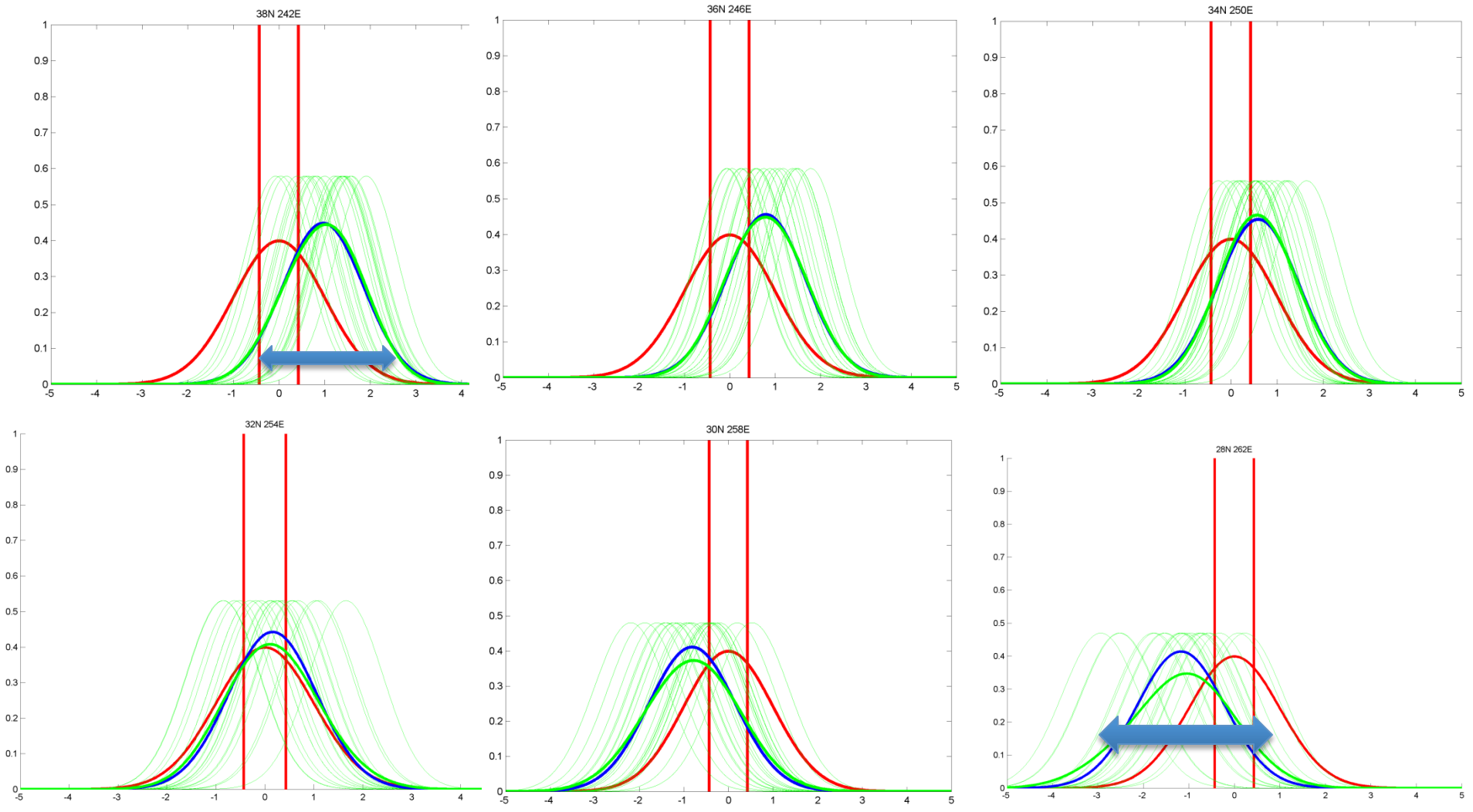
$$N_t = \sqrt{\frac{1}{n}\sum \left(x'_t\right)^2}$$

- By these estimates, signal and noise are dependent on the state of the climate system at time, *t*. (Predictability may be dependent on ENSO, or MJO state, or ?)

- Kernel dressing calibration of ensemble member forecasts (Unger et al., 2009);
- **Shifts in the probability distribution indicate the strength of the signal**



- Shifts in the probability distribution indicate the strength of the signal
- **Spread of ensemble members determine the uncertainty or noise**



Using Ensembles for Predictability Estimates(2)

- A measure of the predictability of the future climate state specific to a subset of initial conditions is the *signal-to-noise ratio*

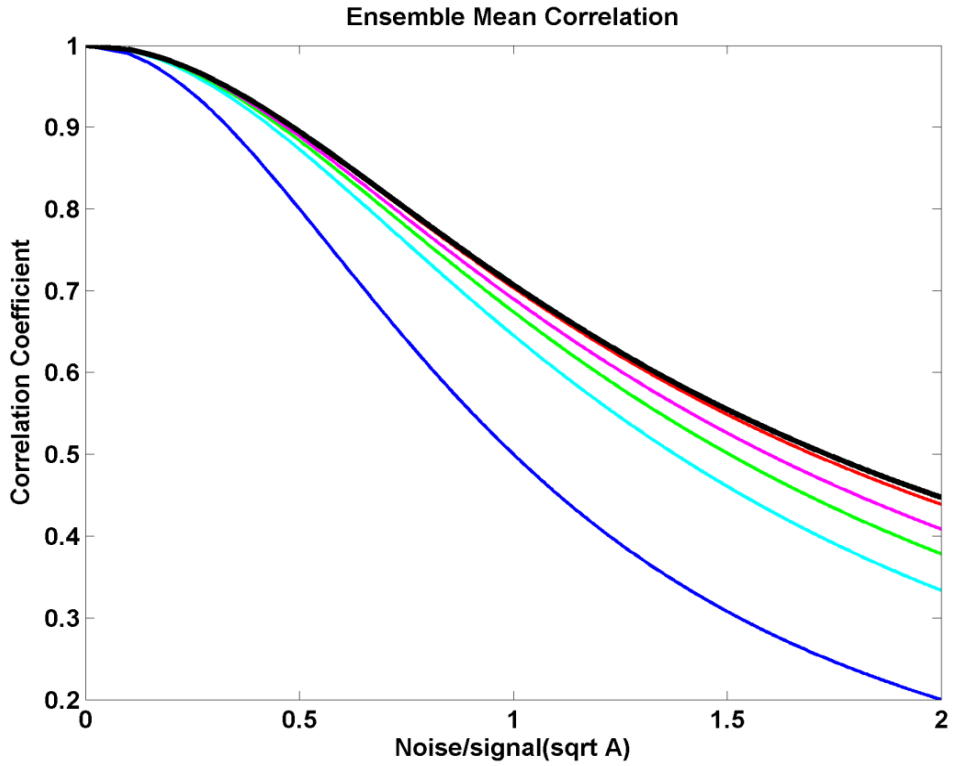
$$SNR_{predictable} = \sqrt{\frac{\overline{S}^2}{\overline{N}^2}} \qquad \overline{S}_{predictable} = \sqrt{\sum_t \langle x_t \rangle^2}$$

i.e. The variability of the ensemble mean relative to the variability of the ensemble members about the mean.

- The variance of the ensemble mean and members can be used as an estimate of the potential skill of forecasts. (Variance is additive.)

$$R_{predictable} = \sqrt{\frac{\overline{S}^2}{\overline{S}^2 + \overline{N}^2}}$$

Potential Correlation as a function of the signal-to-noise ratio, and ensemble size (blue:1 member, green: 10; black: infinite ensemble)



- This places an upper limit on predictability as a function of the SNR (*courtesy of Emily Riddle*)

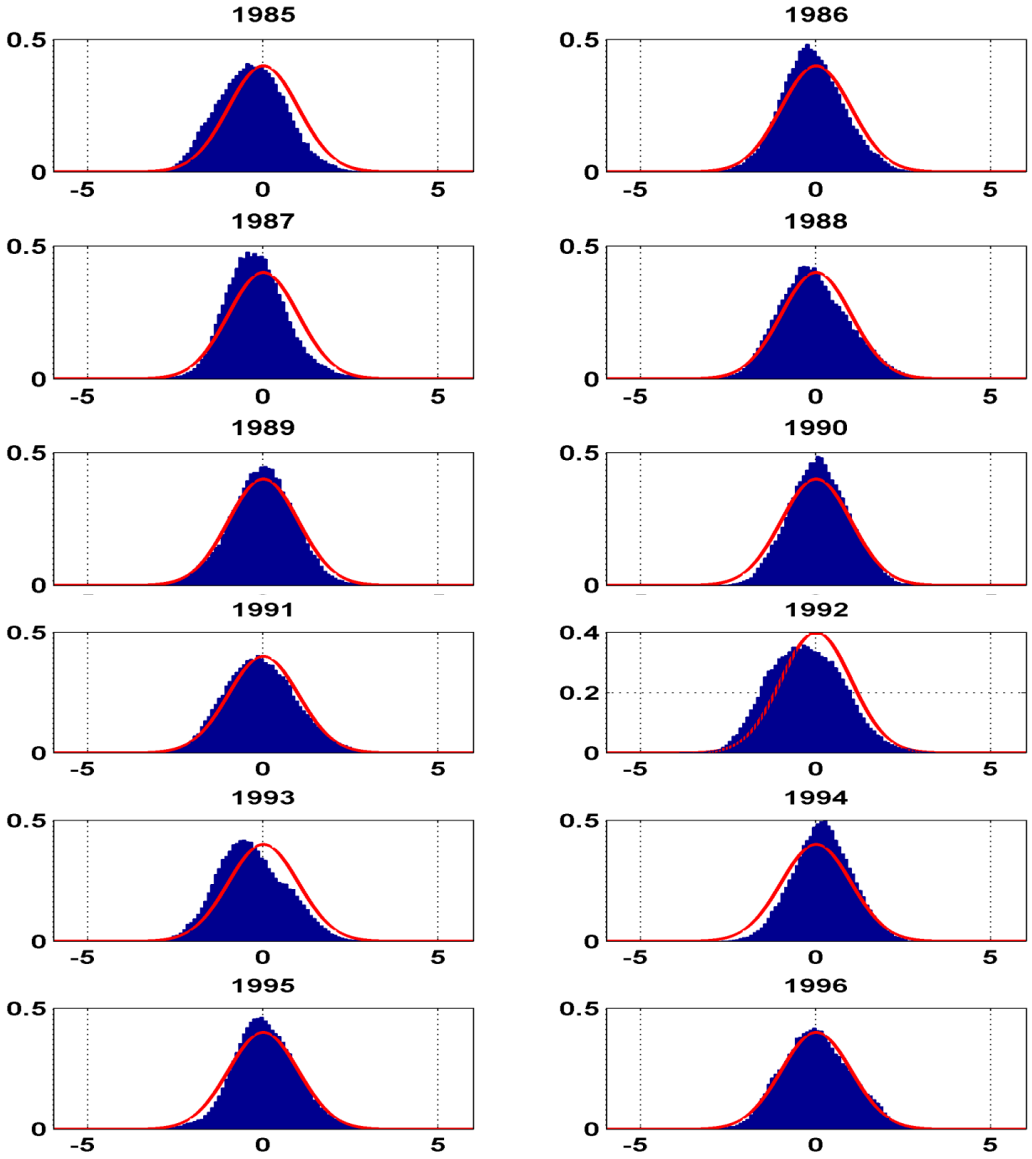
Changes in predictability

- We can consider the linear changes over a number of years

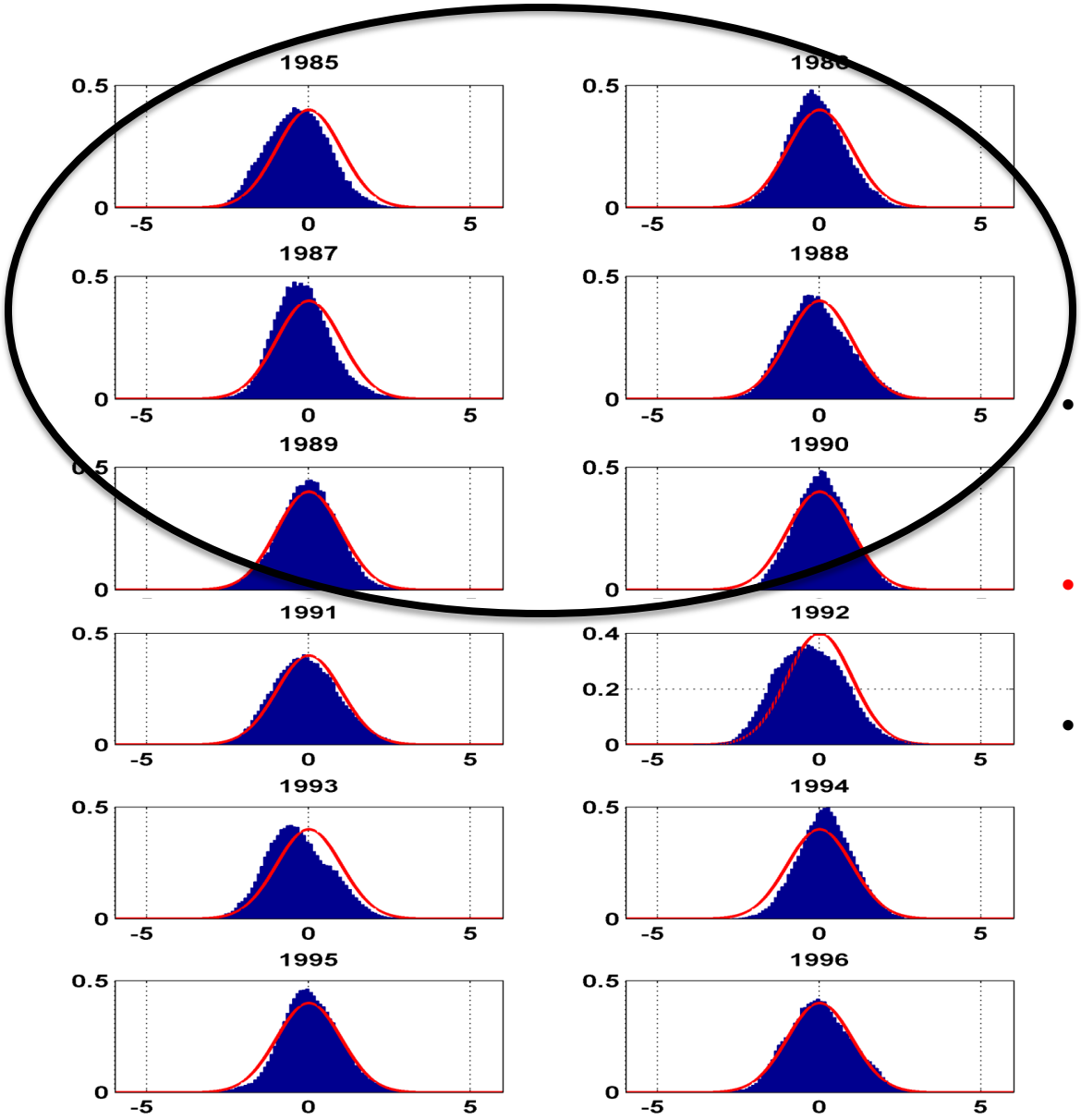
- Linear trends can be calculated for signal (S), noise (N), SNR, and the potential predictability indicated by the prediction of the correlation by the model (R).

It is a reasonable assumption that the predictability of particular characteristics of the climate system may change as the background climate state changes

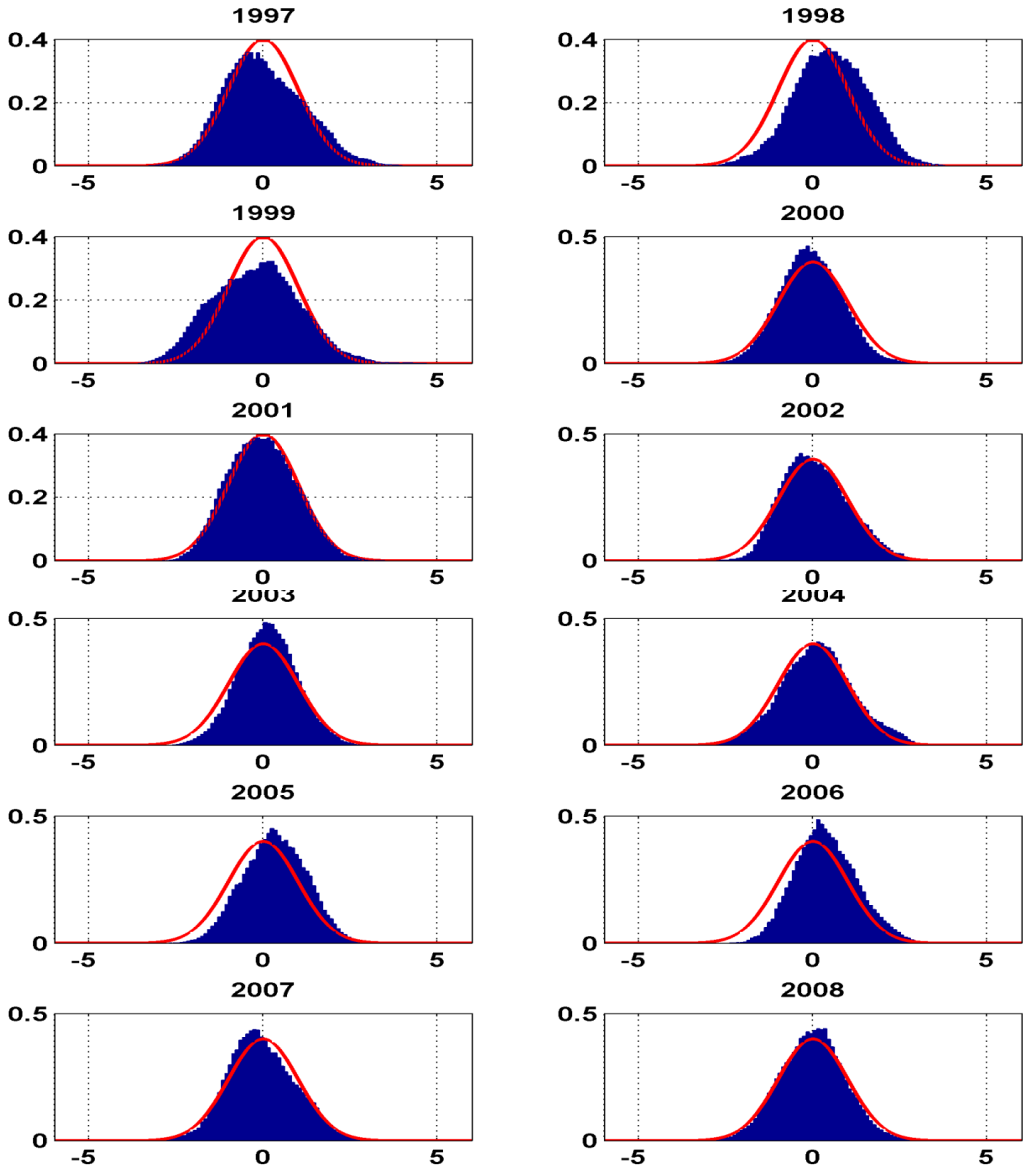
Apparent, predictability may appear to change with changing observational networks.



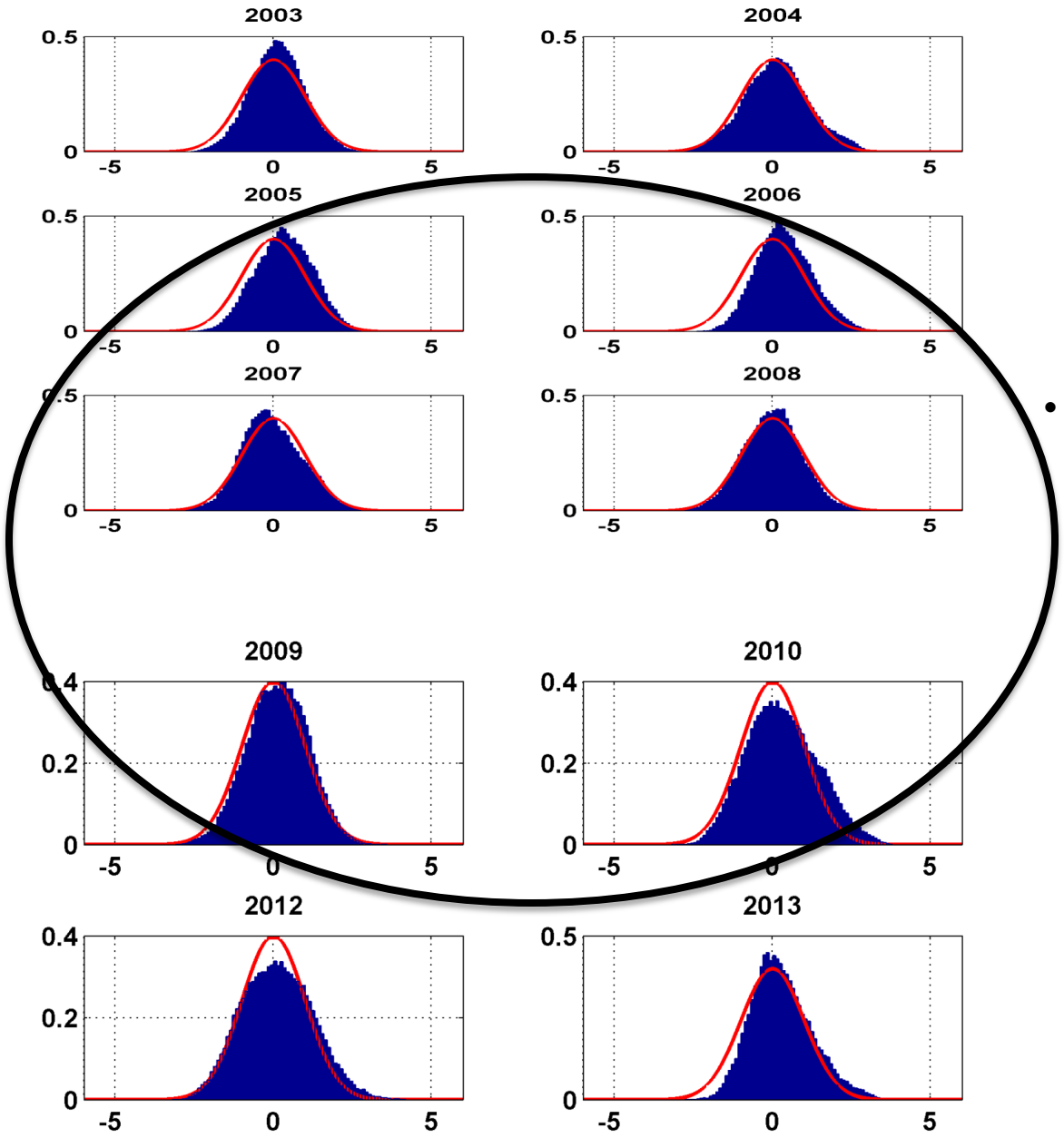
- Weekly mean forecast temperatures over the U.S. by year
- 25-year climatology in red
- PDF shifting each year with interannual variability



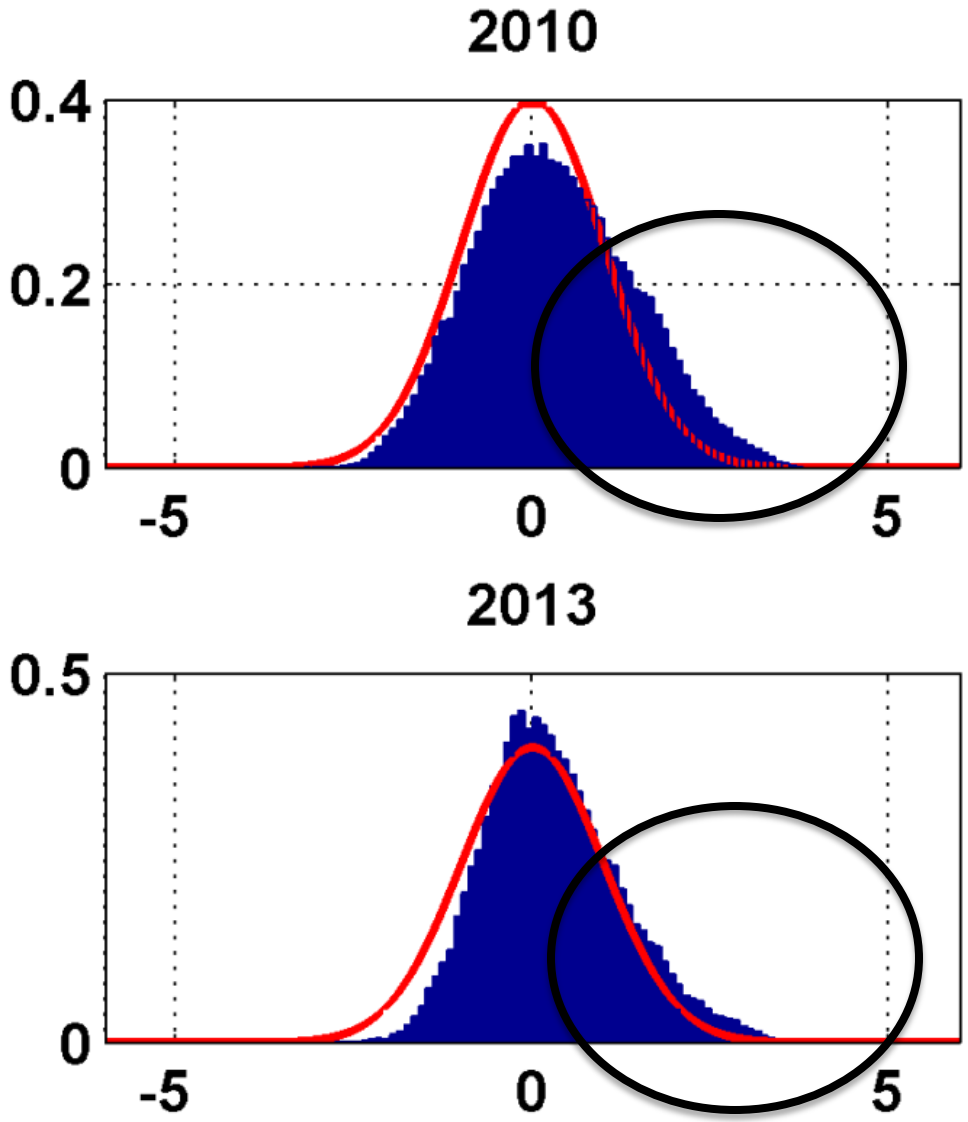
- Weekly mean forecast temperatures over the U.S. by year
- 25-year climatology in red
- PDF shifting each year with interannual variability



- Multi-decadal temperature trends apparent in intraseasonal forecasts



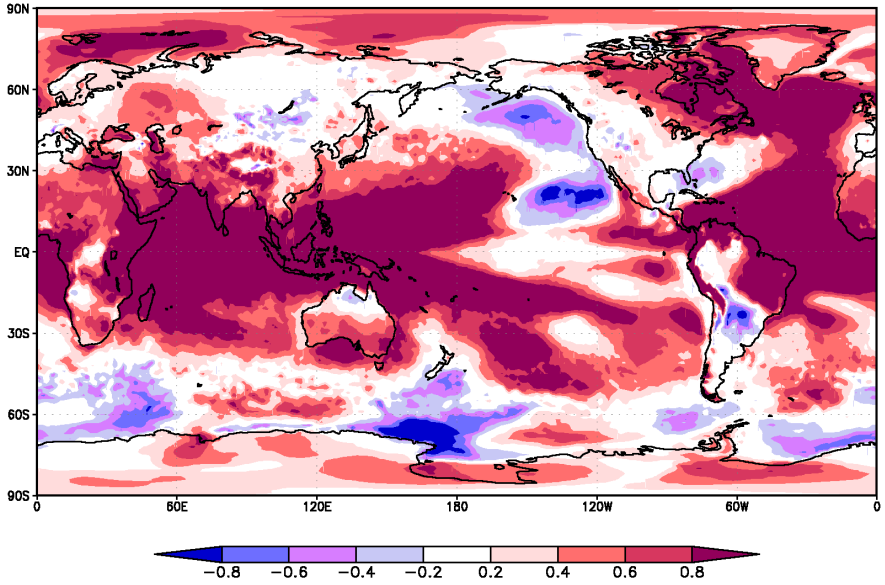
- Shift towards warmer temperatures in more recent years



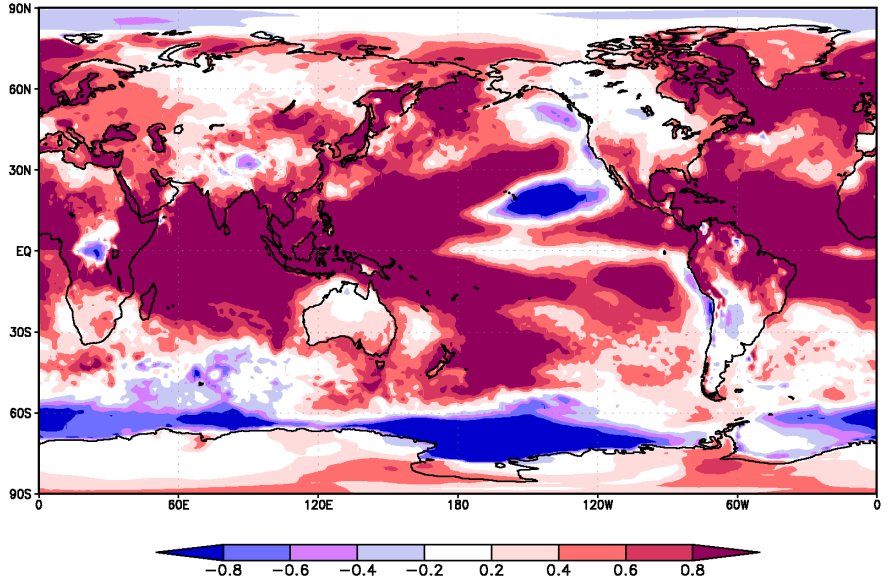
- Increase in the frequency of extremes

CFSR standardized (weekly) linear temperature trend for 25-year period (1985 to 2010)

November to March



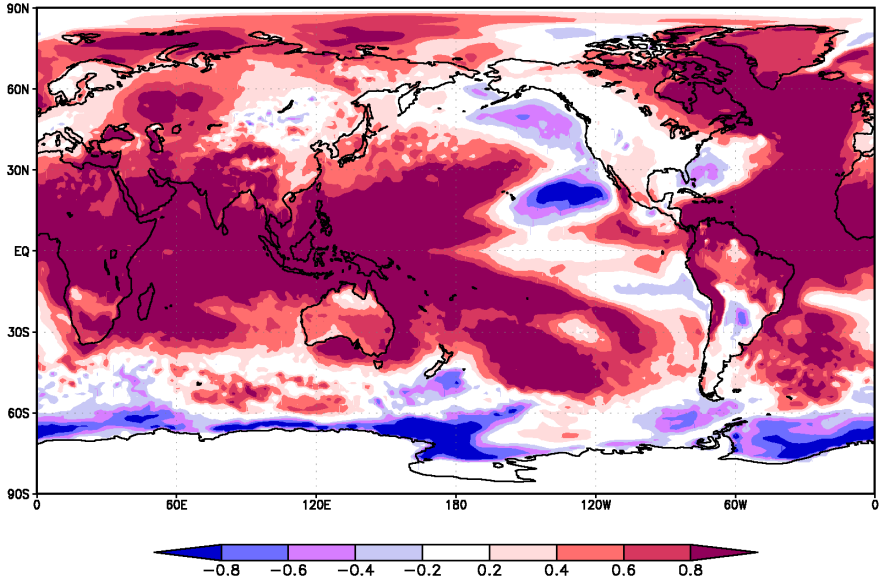
May to September



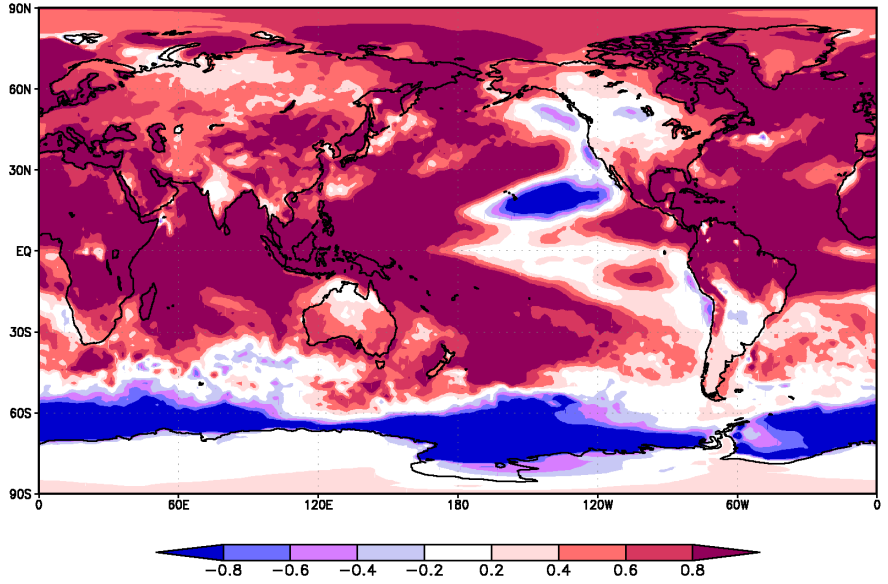
- Temperature trends are significant fraction of weekly timescale variability
- Large areas exceed 0.5 standard deviations

Model week-2 standardized linear temperature trend for 25-year period (1985 to 2010)

November to March



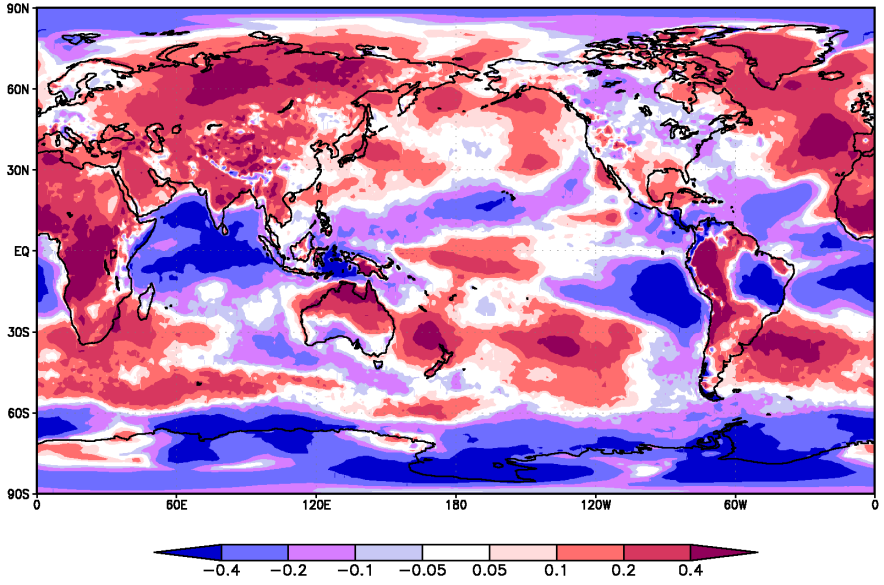
May to September



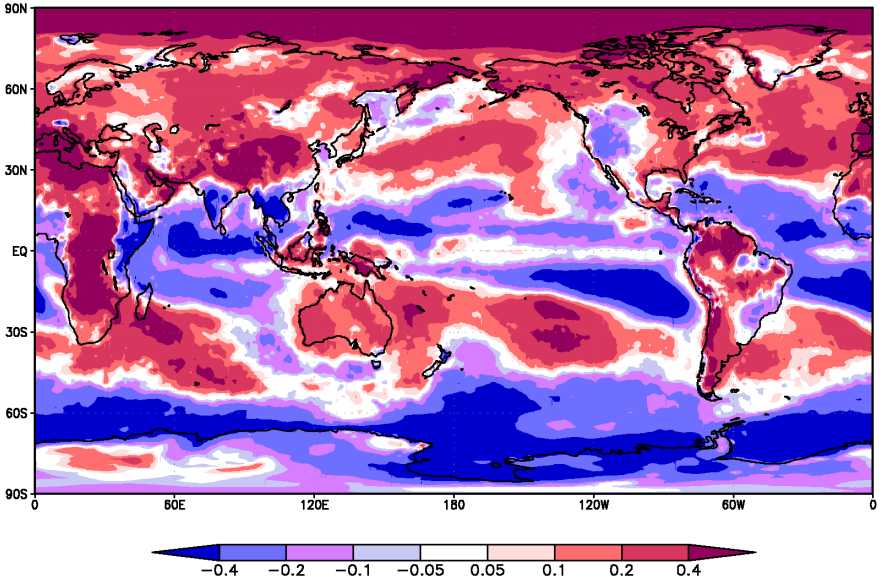
- Decadal changes in the model climate are similar to initialization reanalysis
- Many significant differences globally (bias in the trends)

Standardized trends in **bias** of model week-2 forecasts for 25-year period (1985 to 2010)

November to March



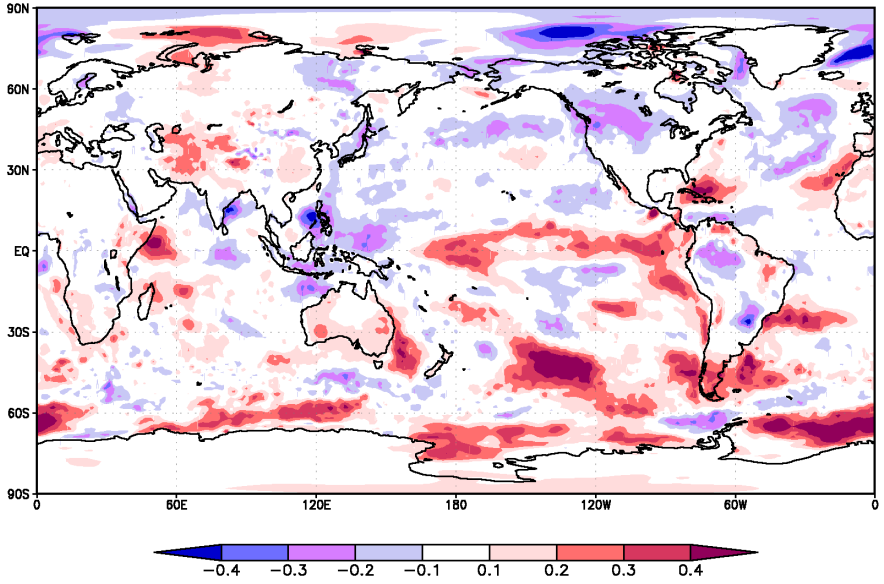
May to September



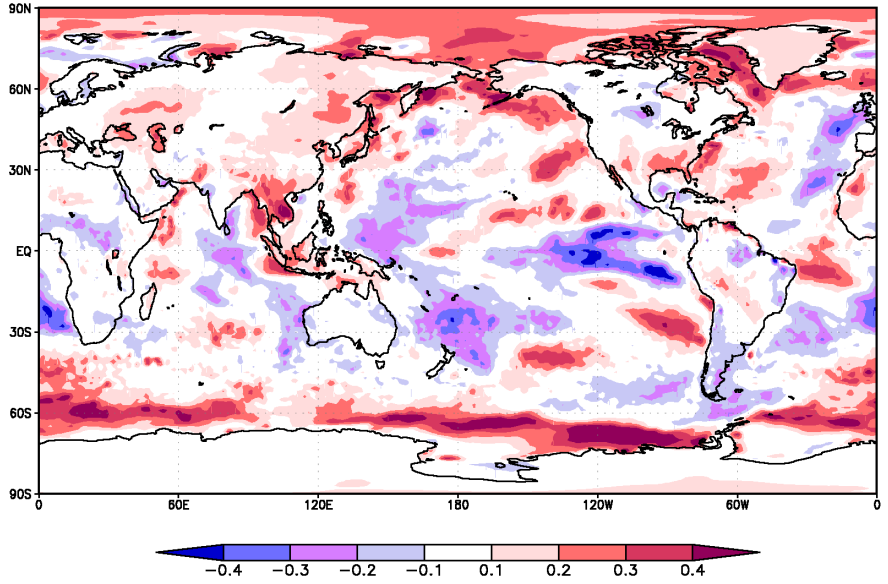
- Bias changes are significant fraction of variability
- Changes in bias over time are highly variable by location and season

Model week-2 standardized trend in **signal** for 25-year period (1985 to 2010)

November to March



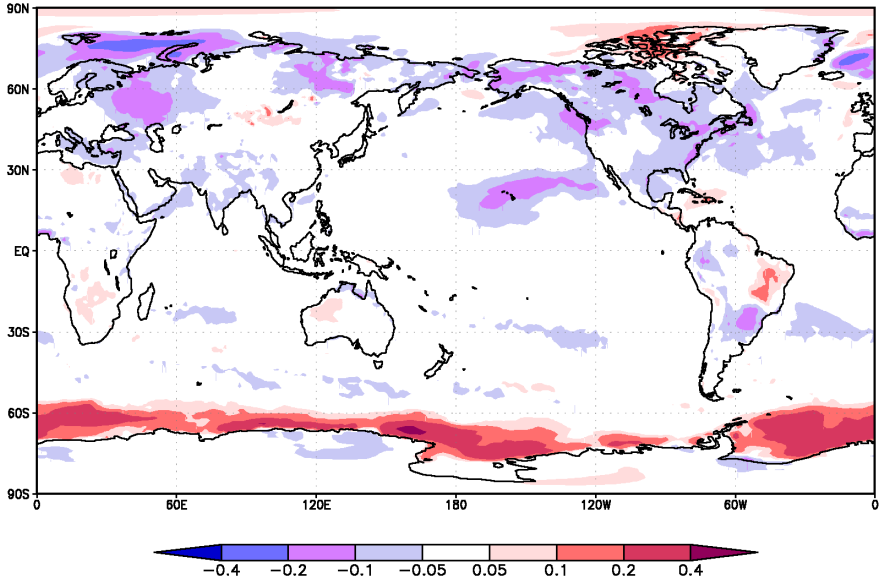
May to September



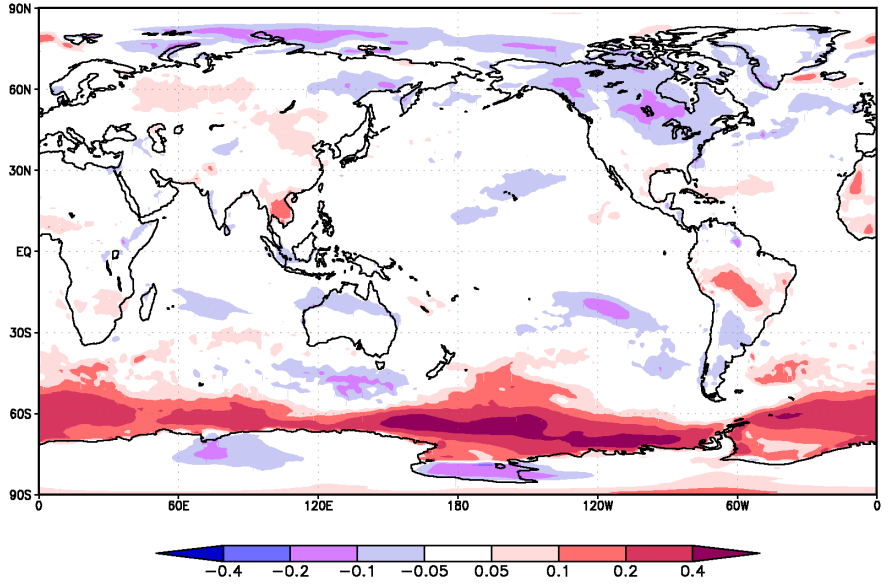
- Decadal changes in signal strength are variable by location and season
- Is this changes in observations? Decadal variability of the climate system? Chance changes in climate variability?

Model week-2 standardized trend in **noise** for 25-year period (1985 to 2010)

November to March



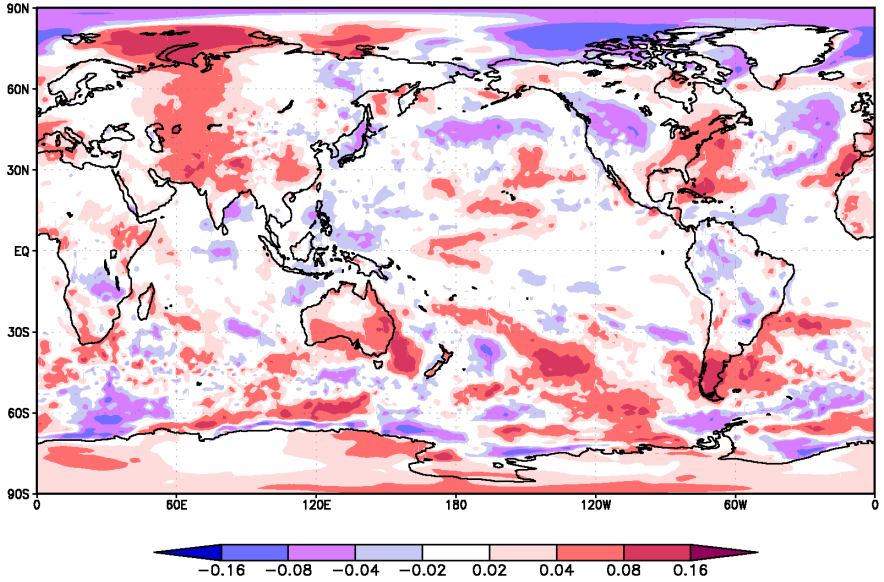
May to September



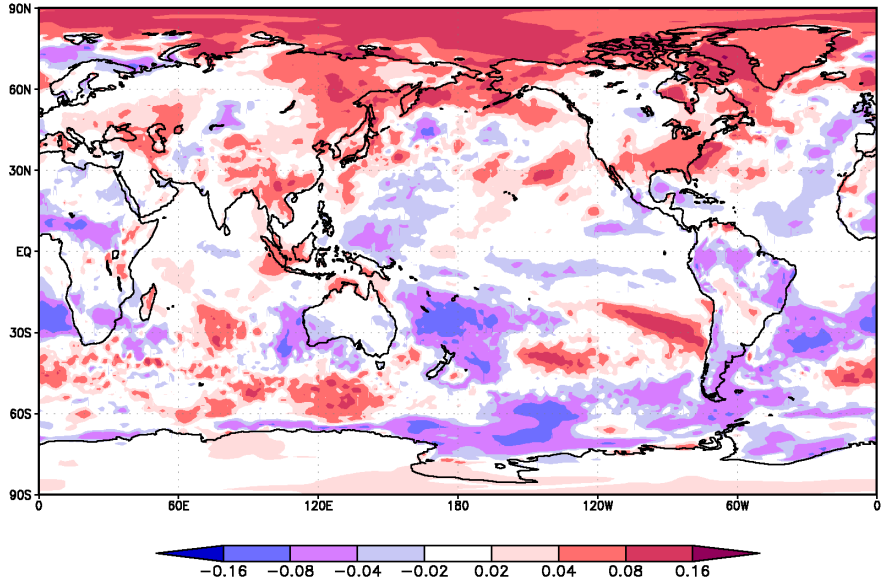
- Negative trend in the climate noise in North America. (Better initial conditions?)
- Weak trends in noise overall

Trend in *model-predicted* **potential correlation (predictability)** for 25-year period (1985 to 2010)

November to March



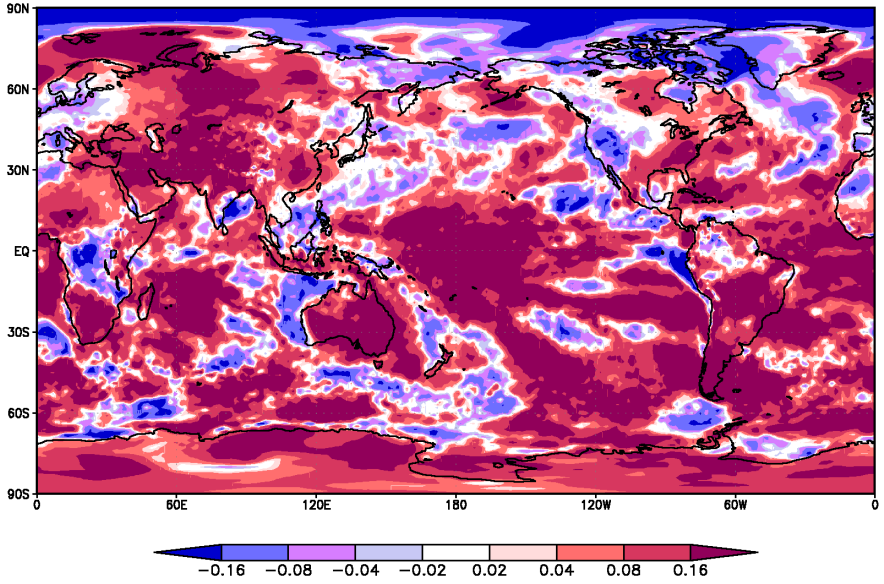
May to September



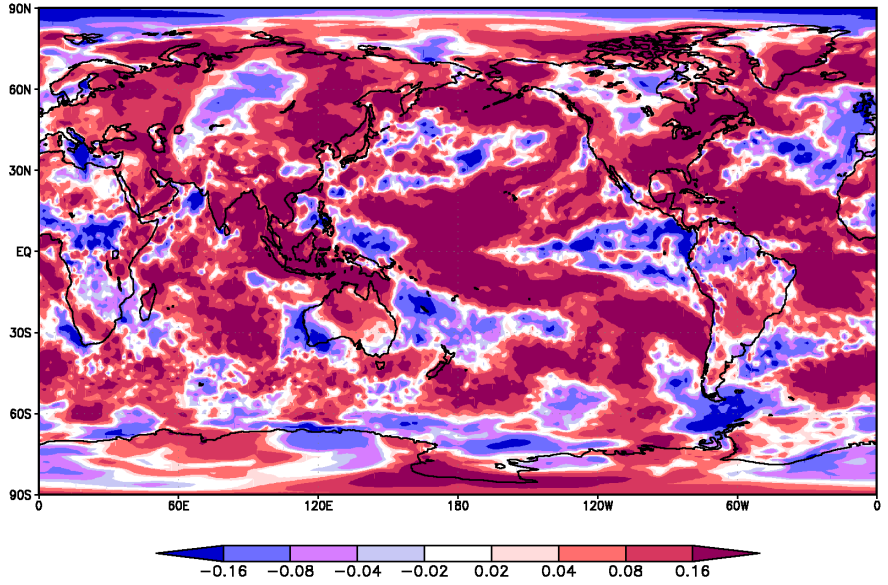
- Potential positive trend in skill in the Summer hemisphere

Model week-2 forecast **observed correlation trend** (actual skill) for 25-year period (1985 to 2010)

November to March



May to September

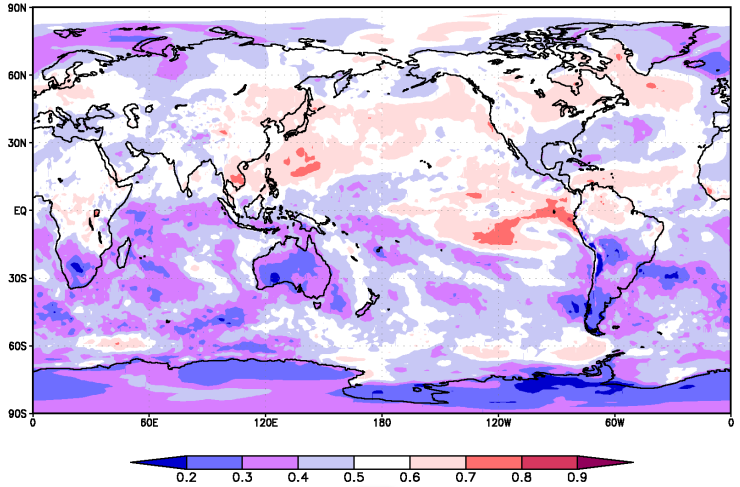


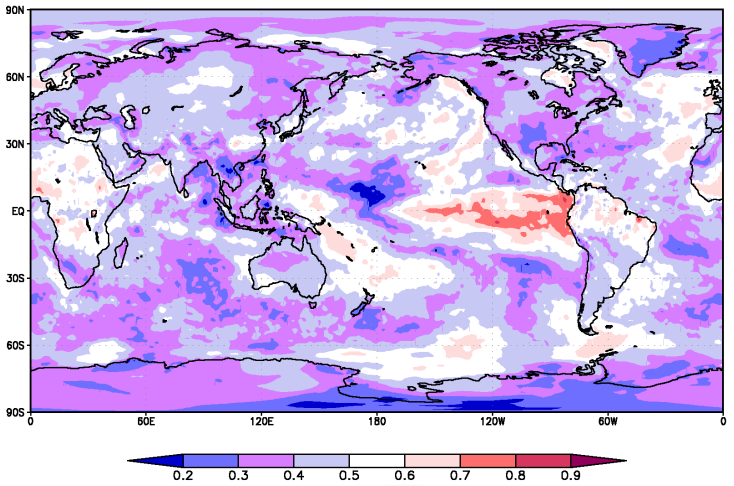
- Some similarities in spatial pattern of trends in correlation to observations compared to the model-only potential correlation trends (previous slide)
- Lower observed correlation relative to potential correlation results in greater increase, as the signal increases in magnitude
- Is this what a better observation network does? Or changing climate-state?

Forecast-analysis **correlation skill** for 1985 and 2010 using the observed trend

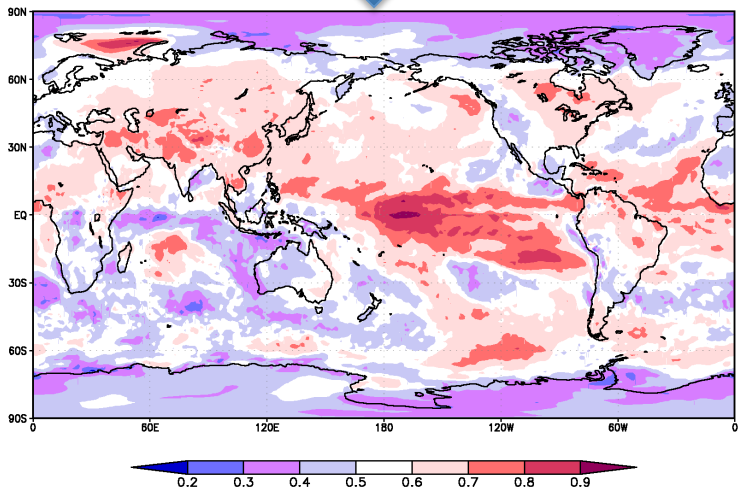
November to March

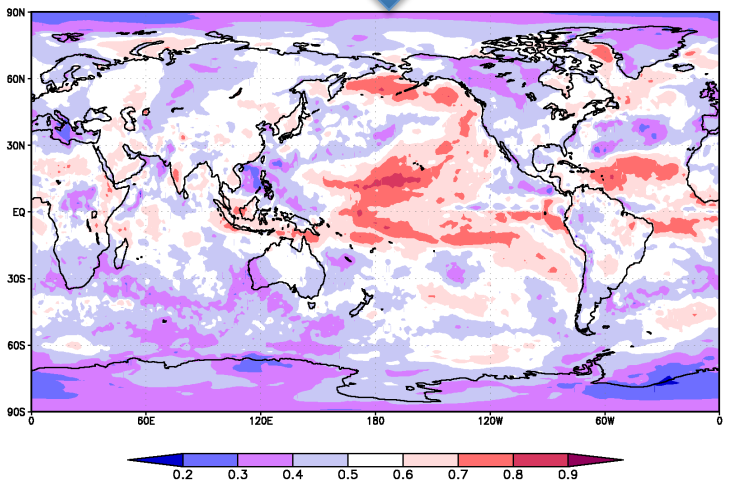
May to September





1985

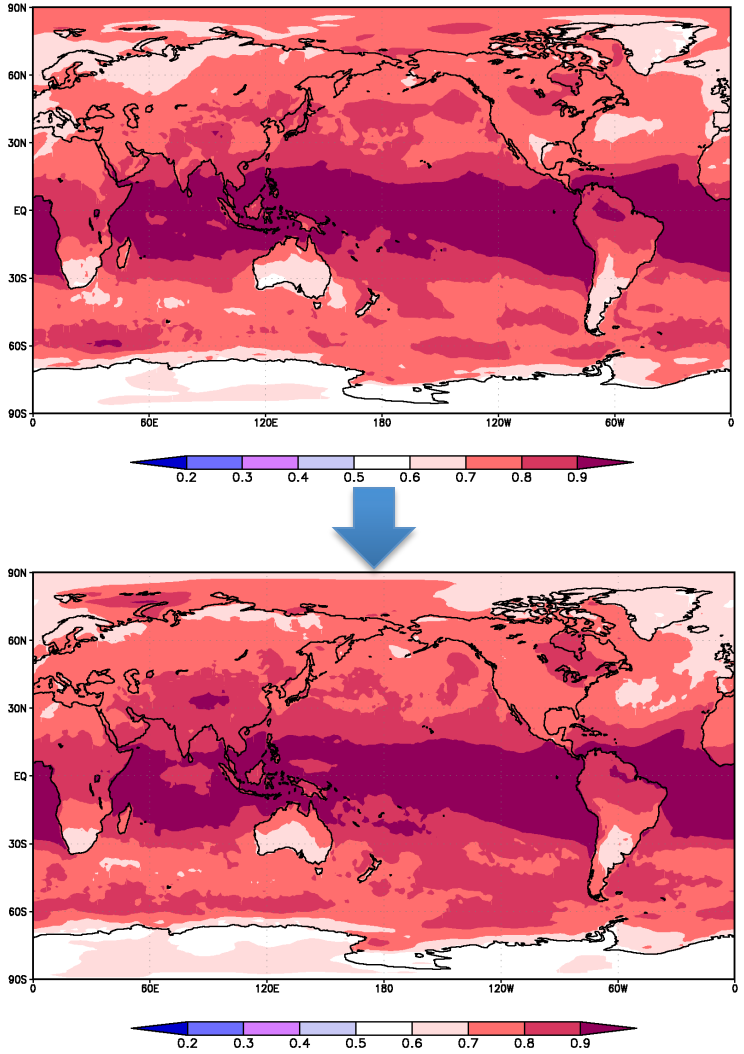




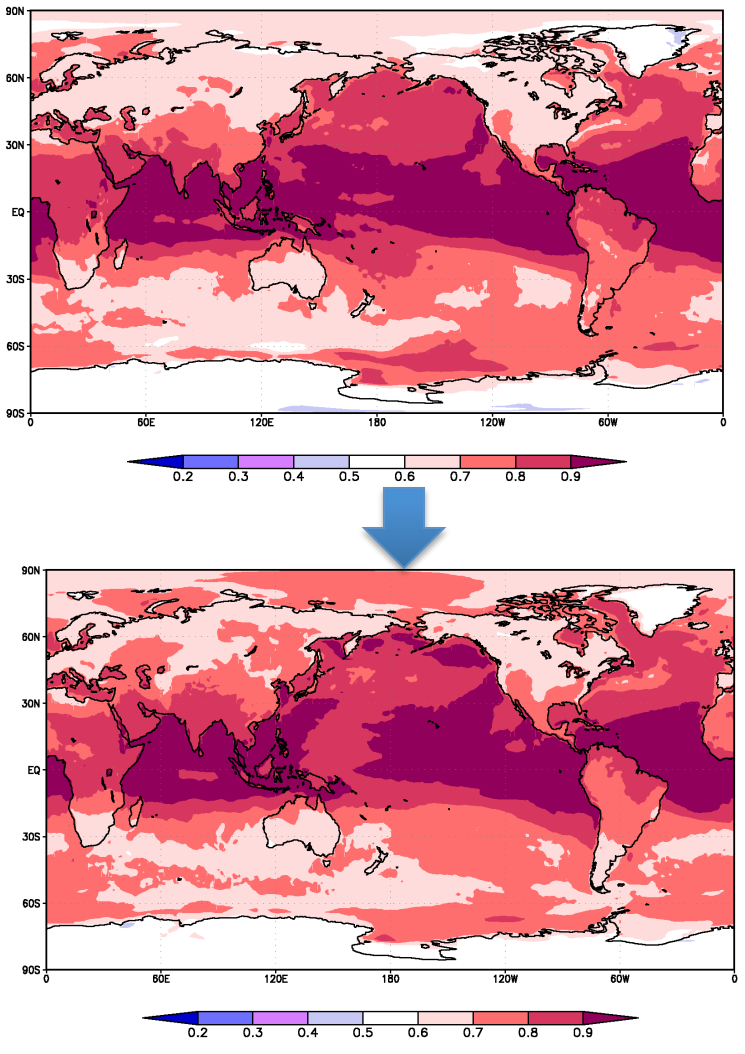
2011

Model week-2 **potential correlation (predictability)** for 1985 (top) and 2010 (bottom) accounting for trend in signal

November to March



May to September

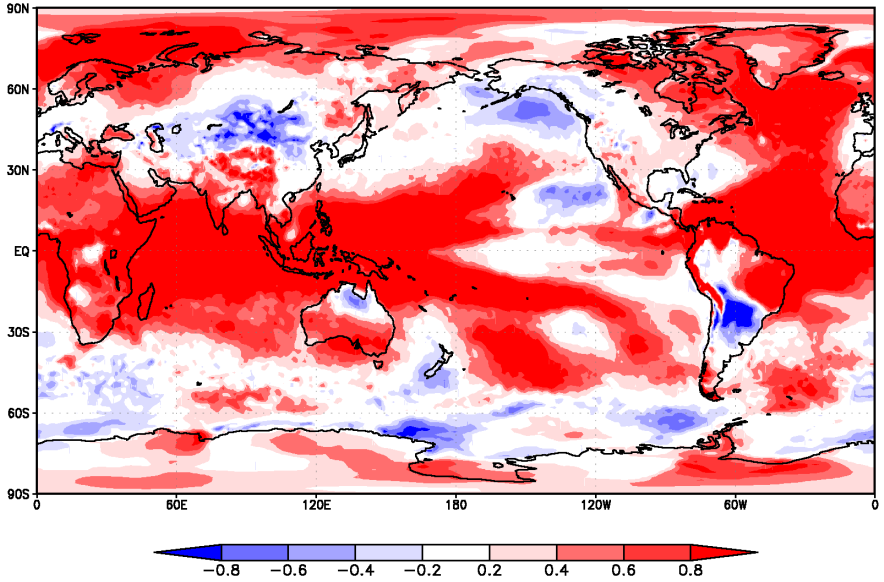


1985

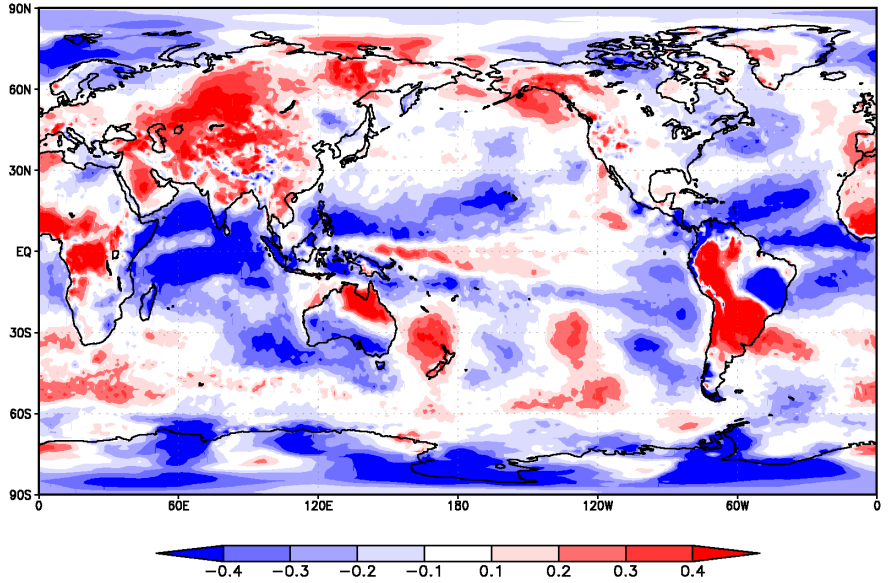
2011

Looking a little more precisely at winter (Dec-Jan) bias and trends

Standardized linear temperature change



Standardized linear trend of ensemble model mean bias



- Growing cold bias where trend is greatest?
- Does greenhouse-gas parameterization (model CO2) matter in week-2 (weather/climate) forecasts?

Summary

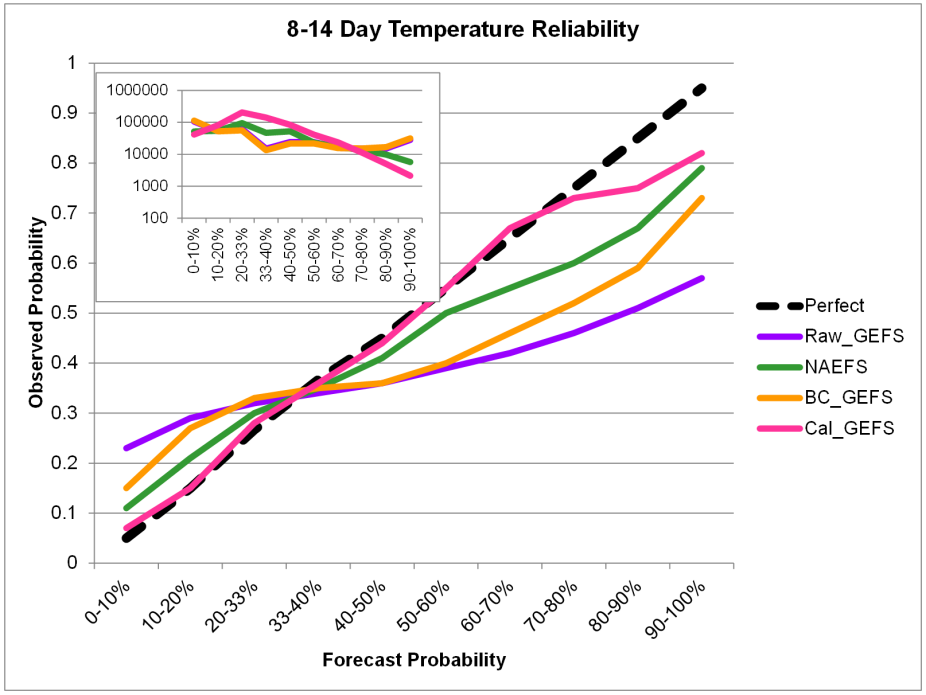
- Decadal temperature trend is significant fraction of subseasonal variance and of predictable signal (in some regions in various seasons)
- Residual variance (noise) after removing the mean variance (signal) is less changeable than signal
- PDF shifts but doesn't change width (noise decreased slightly over N America), which means a change in the signal for forecasts in general and for extremes.
- The systematic bias between the ensemble mean and observations is also changing on decadal and interannual timescales.
- The predictable signal also appears to be non-stationary regionally
  - This may be due to changing observation network and/or changes in climate variability or mean state

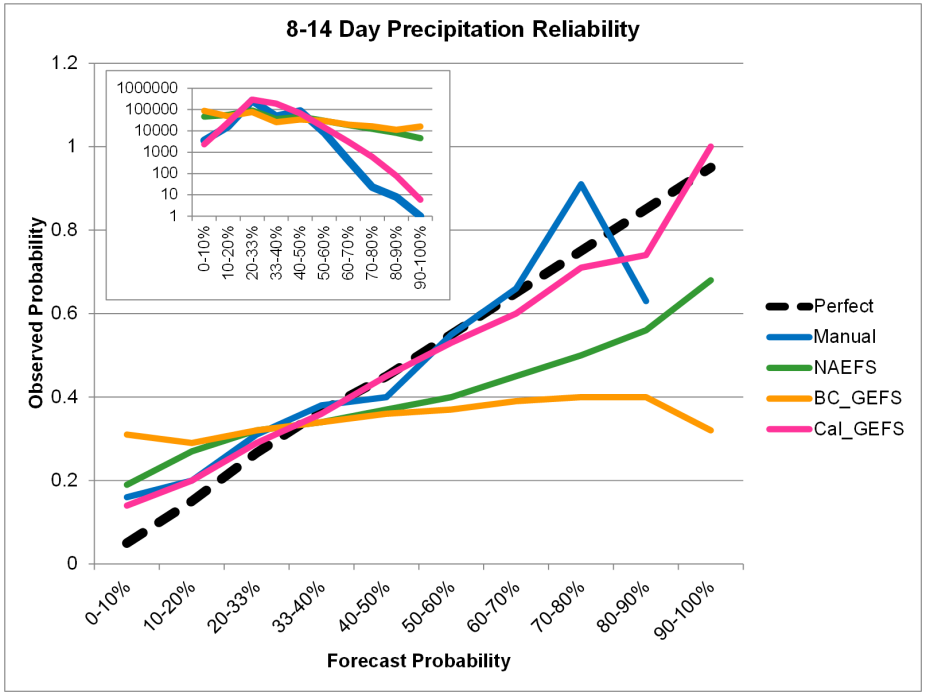
- Many more realizations of the climate in subseasonal forecasts than seasonal
  - Easier to separate timescales of variability in subseasonal vs. seasonal
  - However, all timescales longer than subseasonal are in the average trends
  - Less cases make it more difficult to make similar robust assessments of changing bias and predictability with seasonal forecasts
- Possible non-stationarity of predictability (correlation) of subseasonal variability
  - Why? Better initial conditions; interannual variability (ENSO); climate change

More slides

- More reliability
- Residual bias

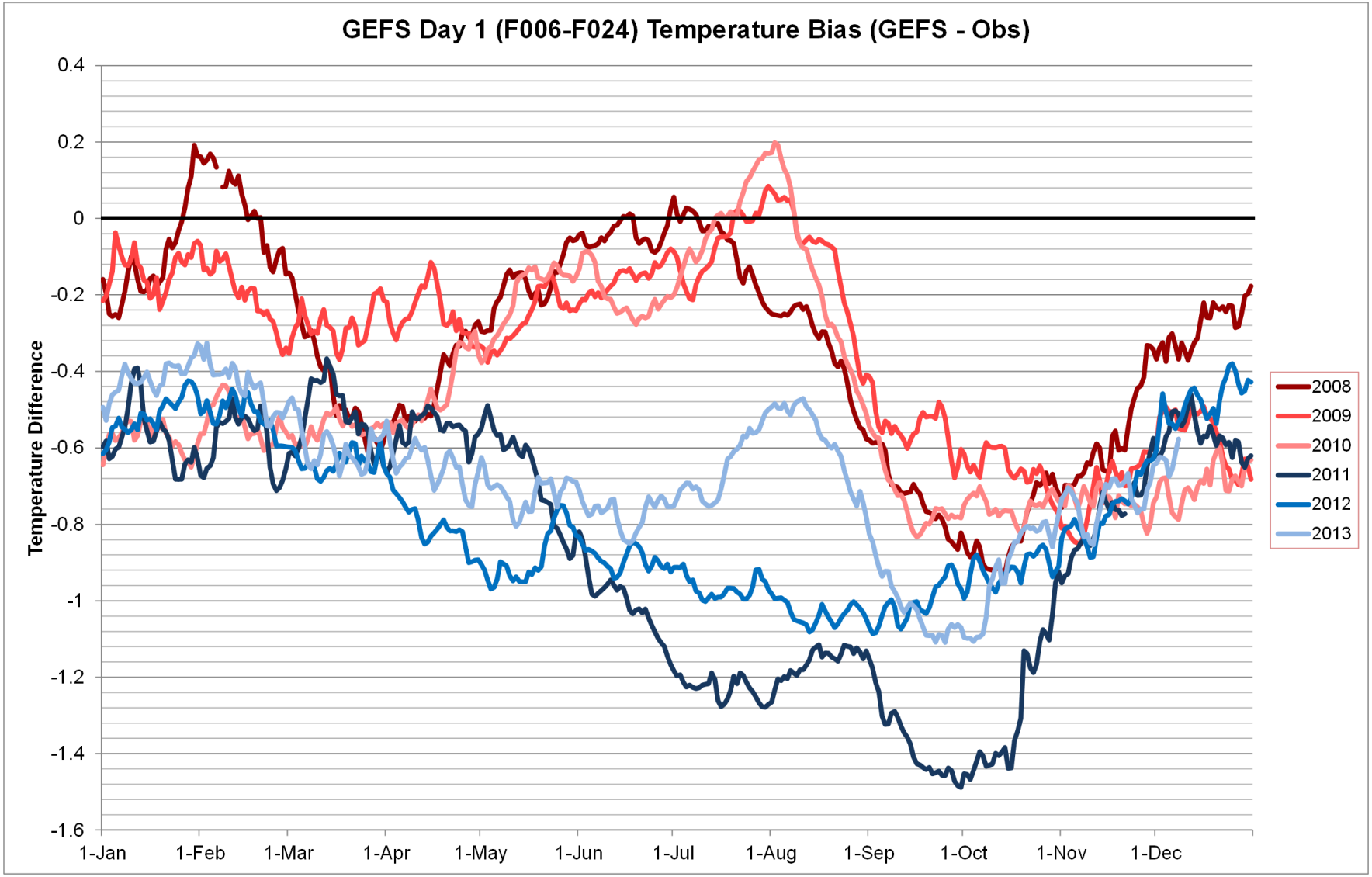
Week-2 reliability of reforecast-calibrated GEFS probabilities compared to NAEFS, bias-corrected GEFS and uncorrected GEFS (3 years 2011-2013)



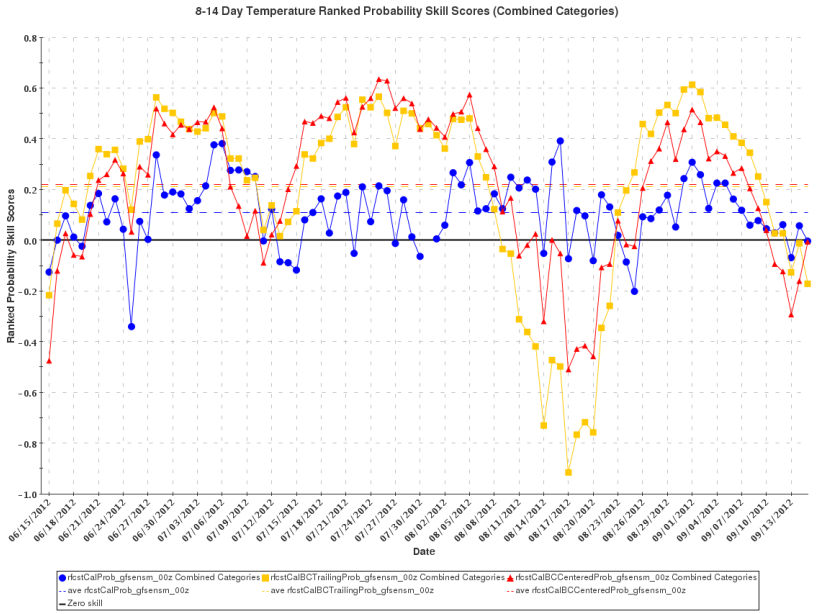


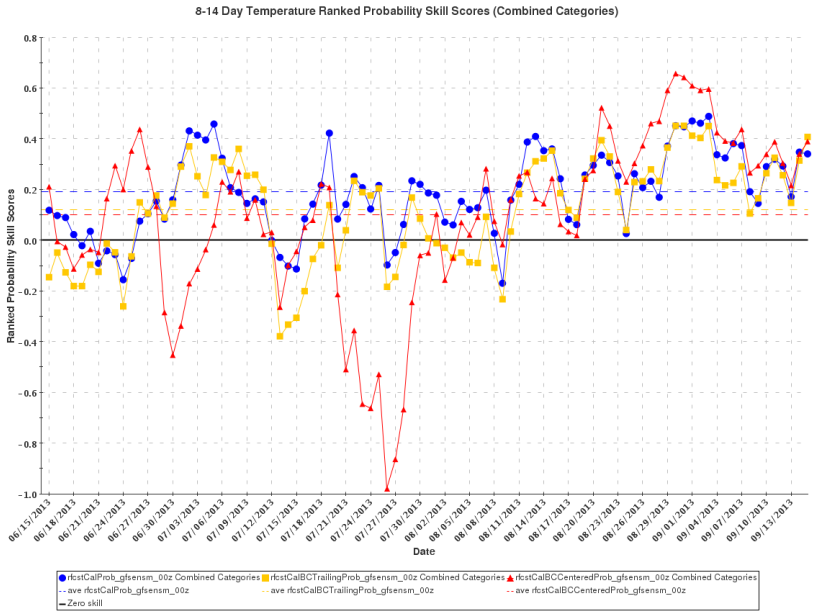
Comparison of reliability of CPC week-2 probability forecasts including raw GEFS, bias-corrected GEFS component of NAEFS, NAEFS MME, and reforecast-calibrated GEFS.

Comparing the time series of model error from the hindcast (red) and real-time forecasts (blue) after bias-correction using hindcast period (1985-2010)



Attempts to correct bias in recent summer forecasts – either trailing period (yellow) or from centered period in real-time forecast years (red) – can have good and bad results relative to using the full hindcast alone





2012: Subtracting residual bias improves RPSS skill

2013: Subtracting residual bias decreases RPSS skill