

# Identifying the capacity of dynamical models to forecast subseasonal extremes: Multi-model ensembles

---

**Dan Collins, Sarah Strazzo and Emerson LaJoie**  
**NOAA Climate Prediction Center**

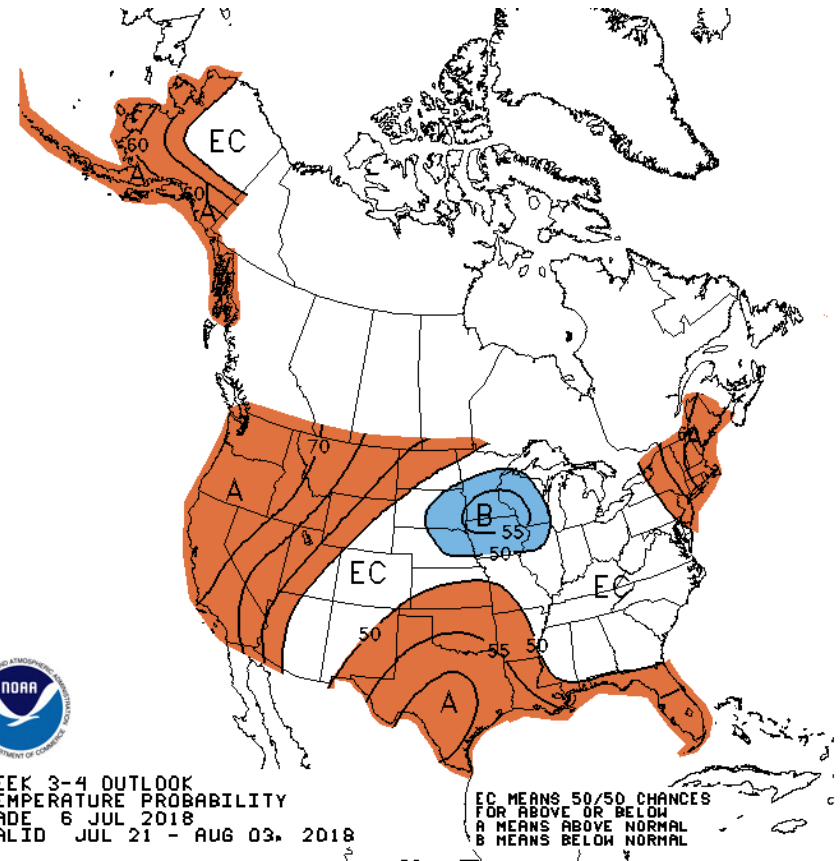
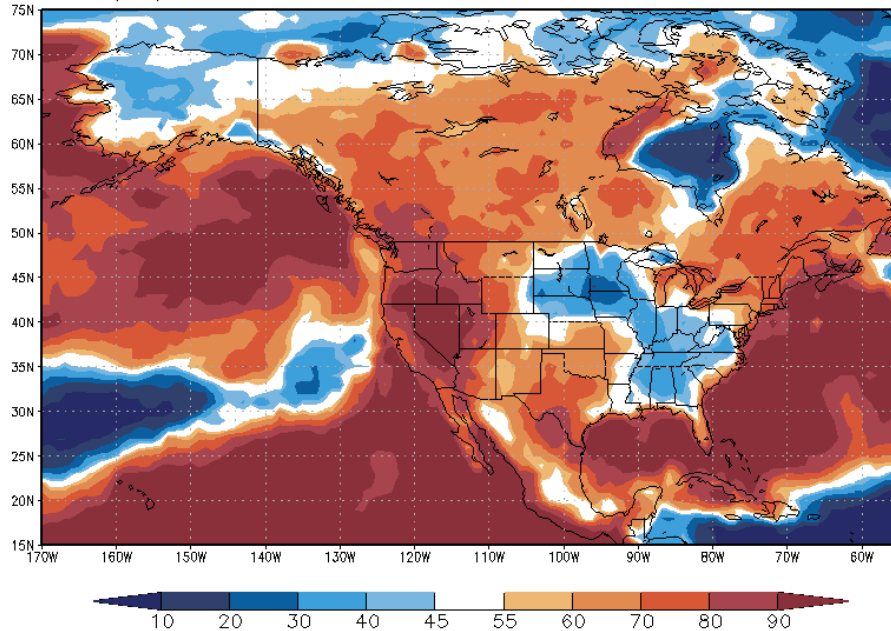
## Outline:

- SubX MME data set
- 3-model mini-MME
- Week 3-4 mean hindcasts, 1999-2014
- Above/below normal (2-category) temperature
- Skill of extreme forecasts (15<sup>th</sup> and 85<sup>th</sup> percentile)
  - ... from climatological distribution
- Observations from CPC global gridded station temperature

# SubX MME real-time probability forecast and NOAA/CPC Outlook

July 6, 2018

MME (79) T Prob Issued: 06 Jul 2018 Valid: 21 Jul – 03 Aug



## Forecasting the tails of the distribution...

- Extremes are by definition intermittent
- Much of the utility of forecasts is in extremes
- Skill of week 3-4 timescale is low but non-zero
- Can usable information be provided?

## Forecasts of Opportunity?

- Intermittent skill  
... that can be identified prior to forecast

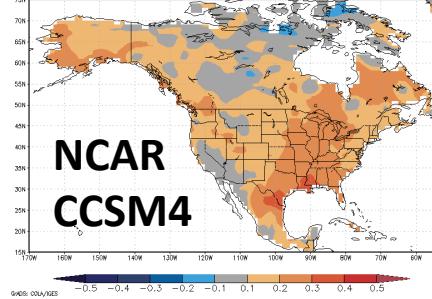


<u>SubX hindcasts</u>	Jan 2	Jan 3	Jan 4	Jan 5	Jan 6	Jan 7	Jan 8	Jan 9 Forecast Day	Week 3-4 Outlook: Jan 24 – Feb 06	
Day of the week & Days to Target Dates	Fri 22:35	Sat 21:34	Sun 20:33	Mon 19:32	Tues 18:31	Wed 17:30	Thurs 16:29	Fri 15:28	2 weeks: Sat + 13 days (Fri) → WK34	
Center-Model ----- Forecast Initialization Period -----										
ECCC-GEM 4 members 32 days								Forecast Day		Real-time forecast day(s)
EMC-GEFS 11 members 35 days								Forecast Day		Variable hindcast days
ESRL-FIMv2 4 members 32 days								Forecast Day		Variable real-time & hindcast days
NASA-GEOS 4 members 45 days								Forecast Day		
NCEP-CFSv2 4 members 44 days								Forecast Day		
NRL-NESM 4 lagged members 45 days								Forecast Day		
RSMAS-CCSM4 3 members 45 days								Forecast Day		
Coming in next year: CESM-46LCAM5 10 members 45 days								Forecast Day		MME for week 3-4 collected for each week in the hindcast
CESM-30LCAM5 10 members 45 days								Forecast Day		

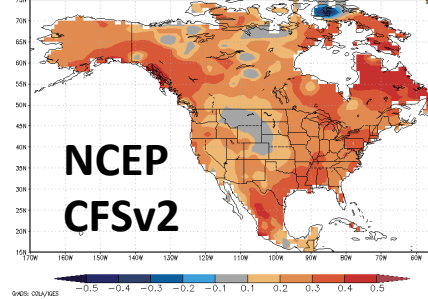
# Anomaly Correlation by model & MME (DJF)

- Preliminary assessments of skill of dynamical model forecasts of above/below normal, for 1999-2014
- **MME outperforms any individual model, as expected**

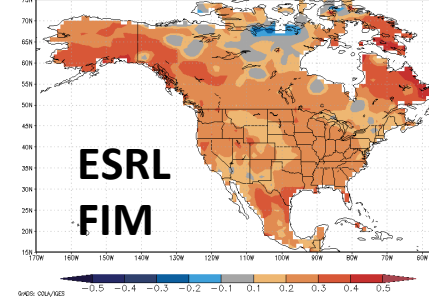
Weighted ACC-DJF TAS RSMAS-CCSM4: Area-avg Score for NA: 0.151



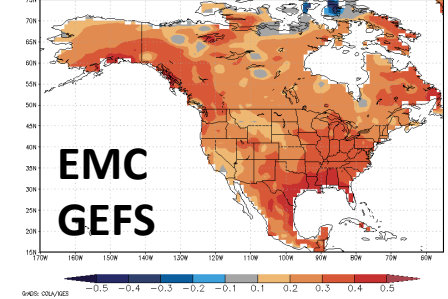
Weighted ACC-DJF TAS NCEP-CFSv2: Area-avg Score for NA: 0.275



Weighted ACC-DJF TAS ESRL-FIMv2: Area-avg Score for NA: 0.254

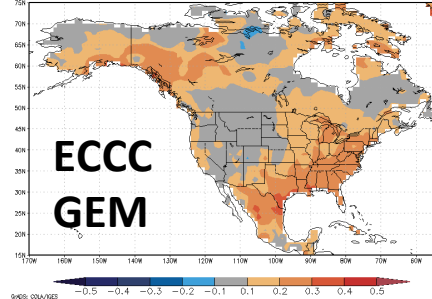


Weighted ACC-DJF TAS EMC-GEFS: Area-avg Score for NA: 0.2753

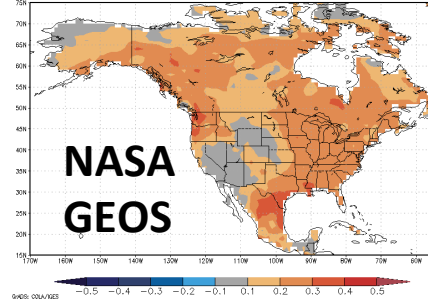


## Individual Model and MME Anomaly Correlations

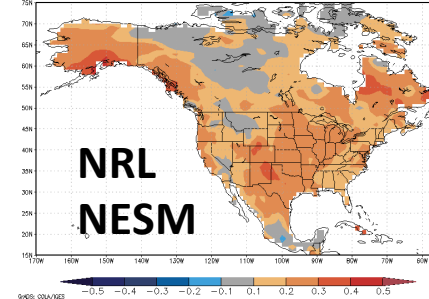
Weighted ACC-DJF TAS ECCC-GEM: Area-avg Score for NA: 0.1380



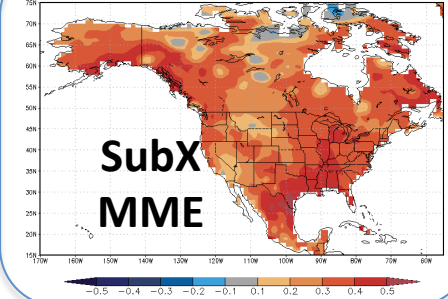
Weighted ACC-DJF TAS NASA-GEOS: Area-avg Score for NA: 0.1916



Weighted ACC-DJF TAS NRL-NESM: Area-avg Score for NA: 0.1931



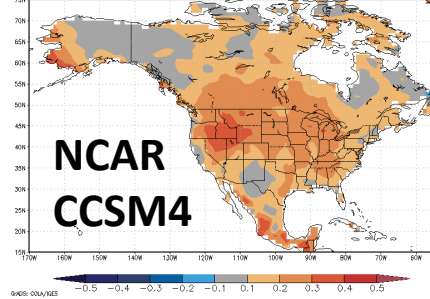
Weighted ACC-DJF TAS 7-MME: Area-avg Score for NA: 0.3252



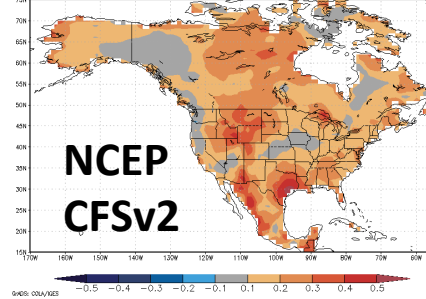
# Anomaly Correlation by model & MME (JJA)

- Preliminary assessments of skill of dynamical model forecasts of above/below normal, for 1999-2014
- **MME outperforms any individual model, as expected**

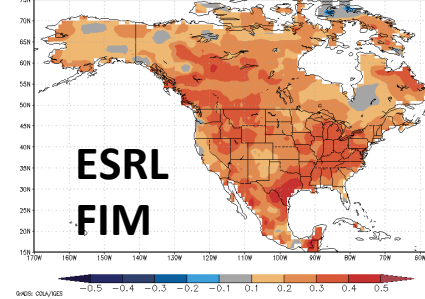
Weighted ACC-JJA TAS RSMAS-CCSM4: Area-avg Score for NA: 0.136



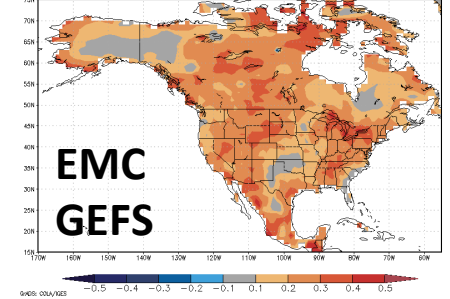
Weighted ACC-JJA TAS NCEP-CFSv2: Area-avg Score for NA: 0.161



Weighted ACC-JJA TAS ESRL-FIMv2: Area-avg Score for NA: 0.2235

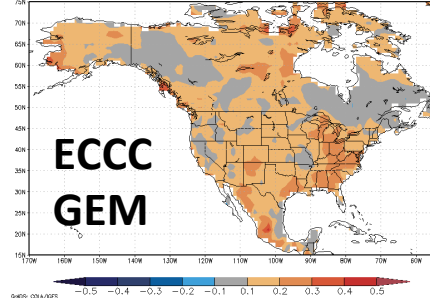


Weighted ACC-JJA TAS EMC-GEFS: Area-avg Score for NA: 0.1948

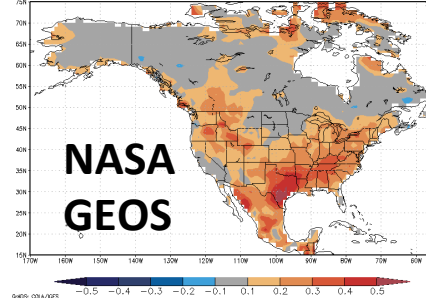


## Individual Model and MME Anomaly Correlations

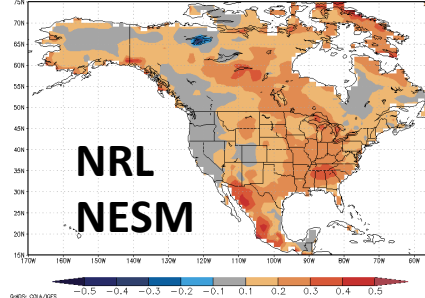
Weighted ACC-JJA TAS ECCC-GEM: Area-avg Score for NA: 0.1244



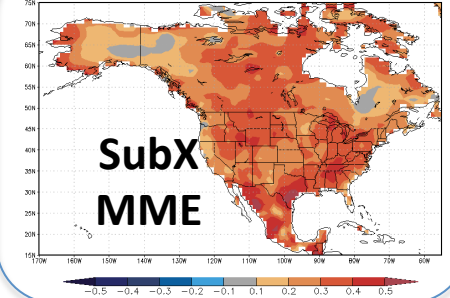
Weighted ACC-JJA TAS NASA-GEOS: Area-avg Score for NA: 0.0906



Weighted ACC-JJA TAS NRL-NESM: Area-avg Score for NA: 0.1531



Weighted ACC-JJA TAS 7-MME: Area-avg Score for NA: 0.2553



<b><u>SubX hindcasts</u></b>	Jan 2	Jan 3	Jan 4	Jan 5	Jan 6	Jan 7	Jan 8	Jan 9 Forecast Day	Week 3-4 Outlook: Jan 24 – Feb 06
<b>Day of the week &amp; Days to Target Dates</b>	Fri 22:35	Sat 21:34	Sun 20:33	Mon 19:32	Tues 18:31	Wed 17:30	Thurs 16:29	Fri 15:28	2 weeks: Sat + 13 days (Fri) → WK34
<b>Center-Model ----- Forecast Initialization Period -----</b>									
ECCC-GEM 4 members 32 days								Forecast Day	Real-time forecast day(s)
EMC-GEFS 11 members 35 days								Forecast Day	Variable hindcast days
ESRL-FIMv2 4 members 32 days								Forecast Day	Variable real-time & hindcast days
NASA-GEOS 4 members 45 days								Forecast Day	
NCEP-CFSv2 4 members 44 days								Forecast Day	
NRL-NESM 4 lagged members 45 days								Forecast Day	
RSMAS-CCSM4 3 members 45 days								Forecast Day	
Coming in next year: CESM-46LCAM5 10 members 45 days								Forecast Day	MME for week 3-4 collected for each week in the hindcast
CESM-30LCAM5 10 members 45 days								Forecast Day	

# SubX mini-MME

- 3 models with Wednesday starts

(Week 3-4 = days 17-30)

- NCEP CFSv2 (4 members)
- ESRL FIM (4 members)
- NCEP/EMC GEFS (11 members)
- 19 members total

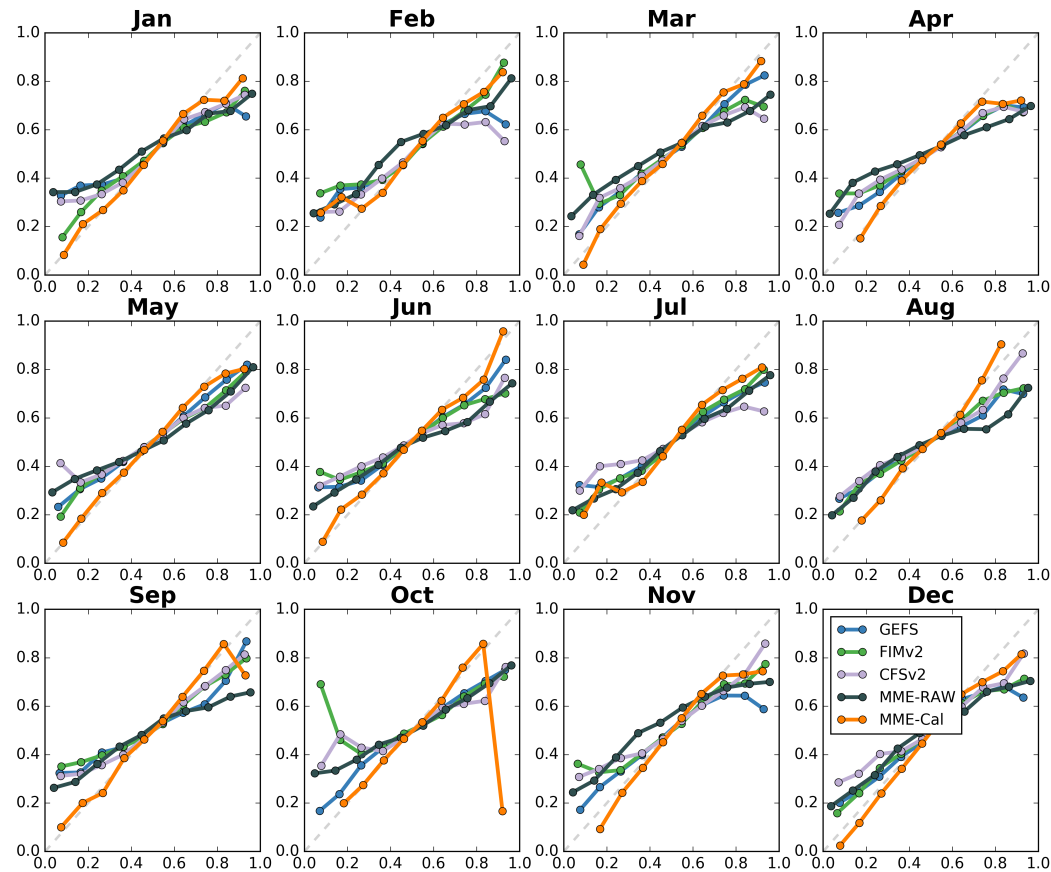
- Calibration of models by Bayesian Joint Probability models (Wang et al. 2009)

Wang, Q., D. Robertson, and F. Chiew, 2009: A Bayesian joint probability modeling approach for seasonal forecasting of streamflows at multiple sites. Water Resources Research, 45 (5).

- Forecasting extreme below & extreme above (15<sup>th</sup> & 85<sup>th</sup> percentiles from observed climatology)

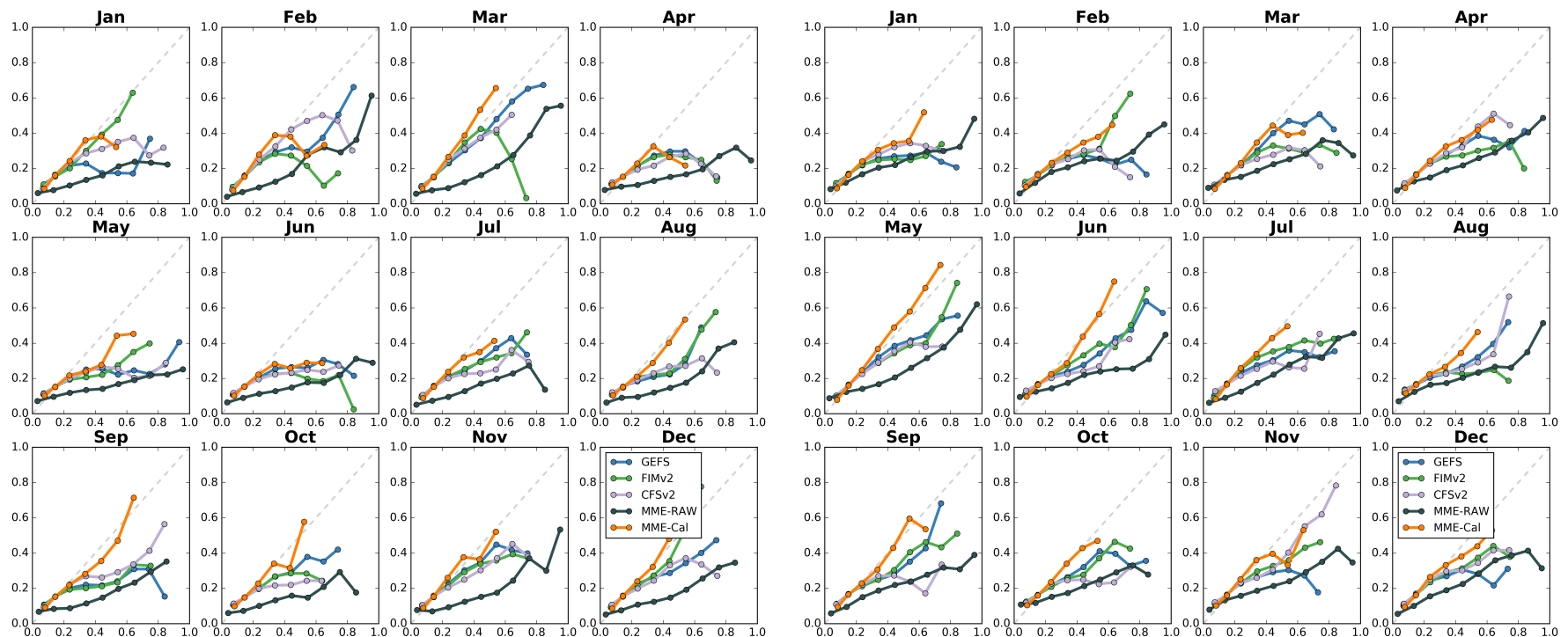
# Above / below normal temperature Reliability

- **Calibrated MME** improves reliability of MME in all months
- **Calibrated MME** more reliable than calibrated **SubXGEFS**, **FIMv2** or **CFSv2**, or **raw MME count** of ensemble members



# Extreme above / below normal reliability

- **Calibrated MME** essential to extreme probability reliability
- **Raw MME** much less reliable probabilities
- Calibrated **SubXGEFS**, **FIMv2** or **CFSv2** less reliable than **MME**

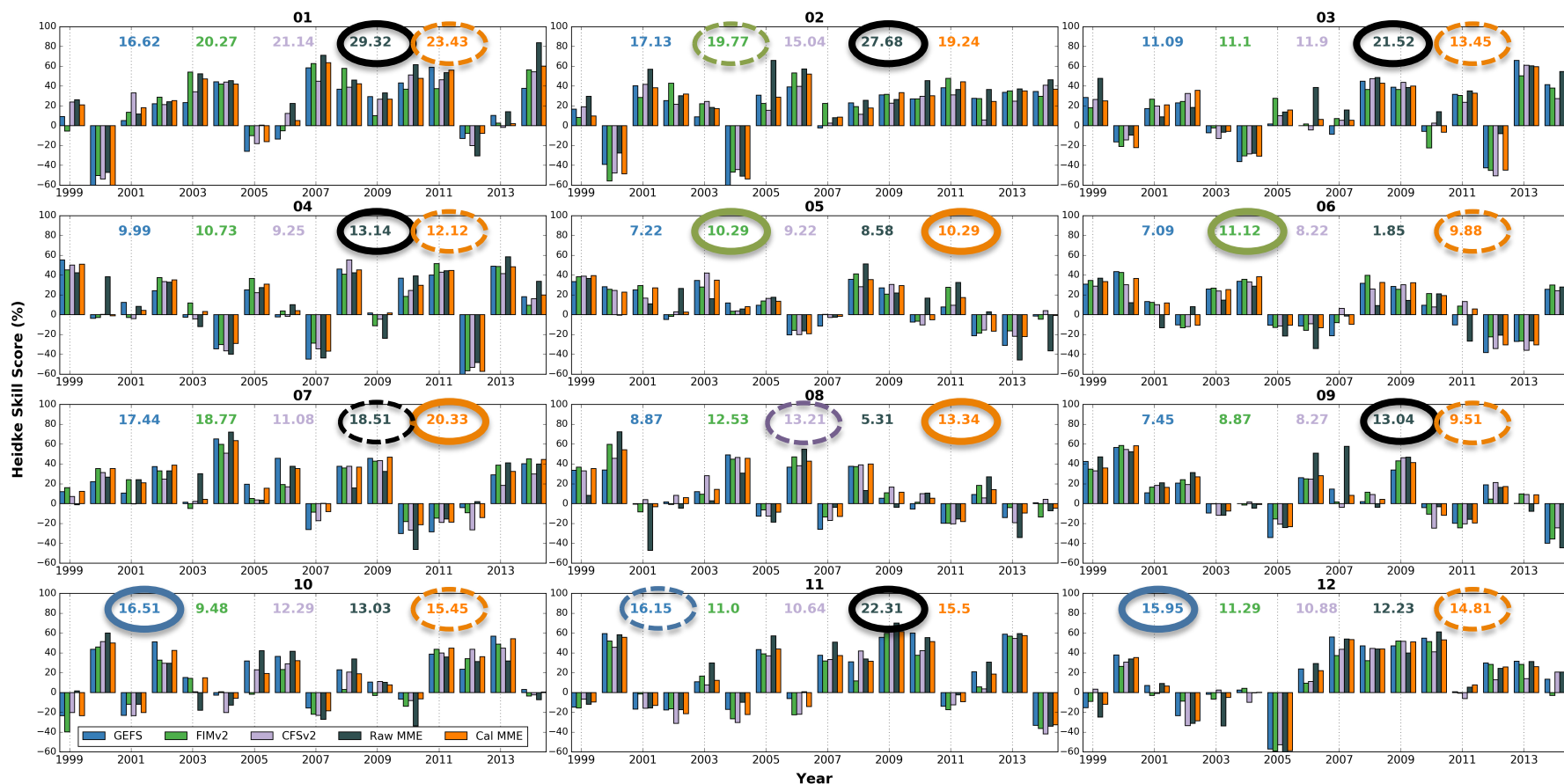




# Monthly (1999-2014) above normal temperature Heidke Skill Score

## 1<sup>st</sup> & 2<sup>nd</sup> ranked models

- **Calibration** of raw mini-MME probabilities *improves overall Heidke Skill Score*
- **Raw mini-MME** has less reliable probabilities but occasionally better hit rate
- MME more skillful in most months / years than **GEFS**, **FIMv2** or **CFSv2**





# Extreme **above normal** temperature Heidke Skill

- **Calibrated mini-MME** improves Heidke Skill Score in all months / most years
- **Raw mini-MME** less reliable probabilities and generally poor hit rate
- **Calibrated** more skillful than **GEFS**, **FIMv2** or **CFSv2** in nearly all months



# Extreme **below normal** temperature Heidke Skill

- **Calibrated MME** improves Heidke Skill Score in all months / most years
- **Raw MME** less reliable probabilities and lower hit rate
- **Calibrated** more skillful than **GEFS**, **FIMv2** or **CFSv2** in nearly all months



# **“Forecasts of Opportunity”**

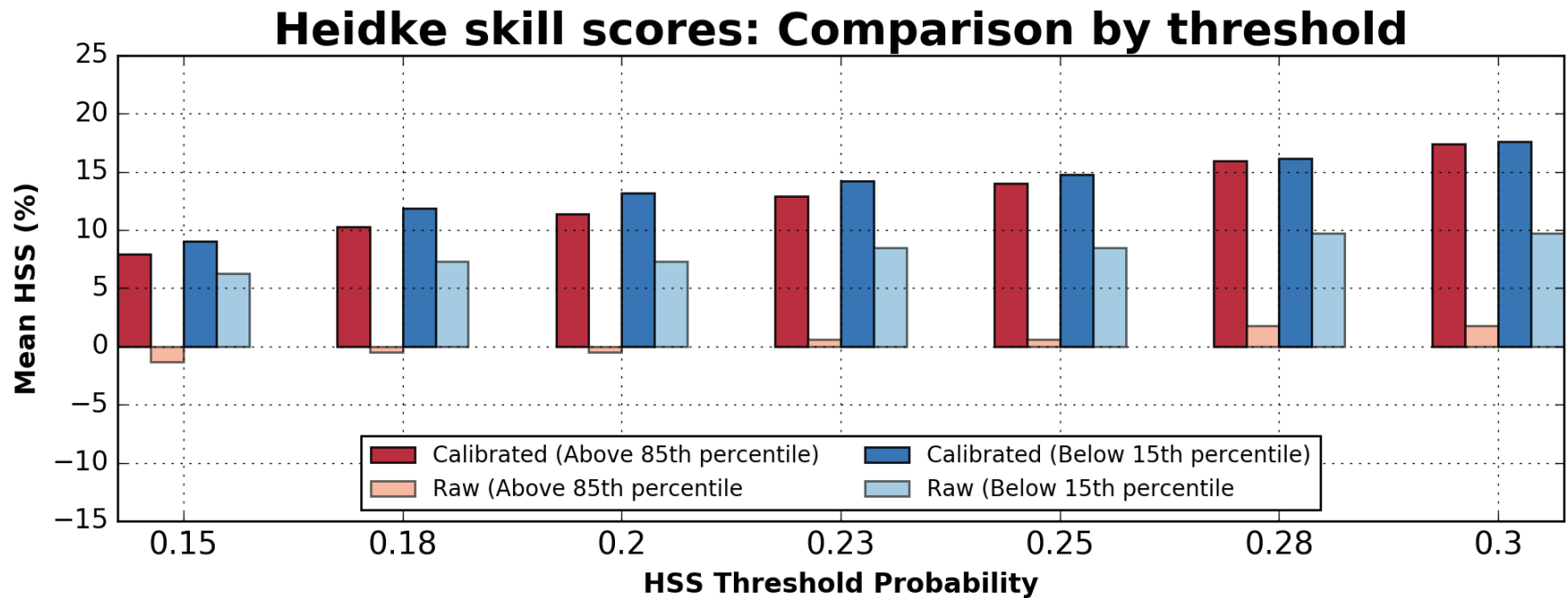


# Intermittent skill of forecasts

... that can be identified **prior to forecast**

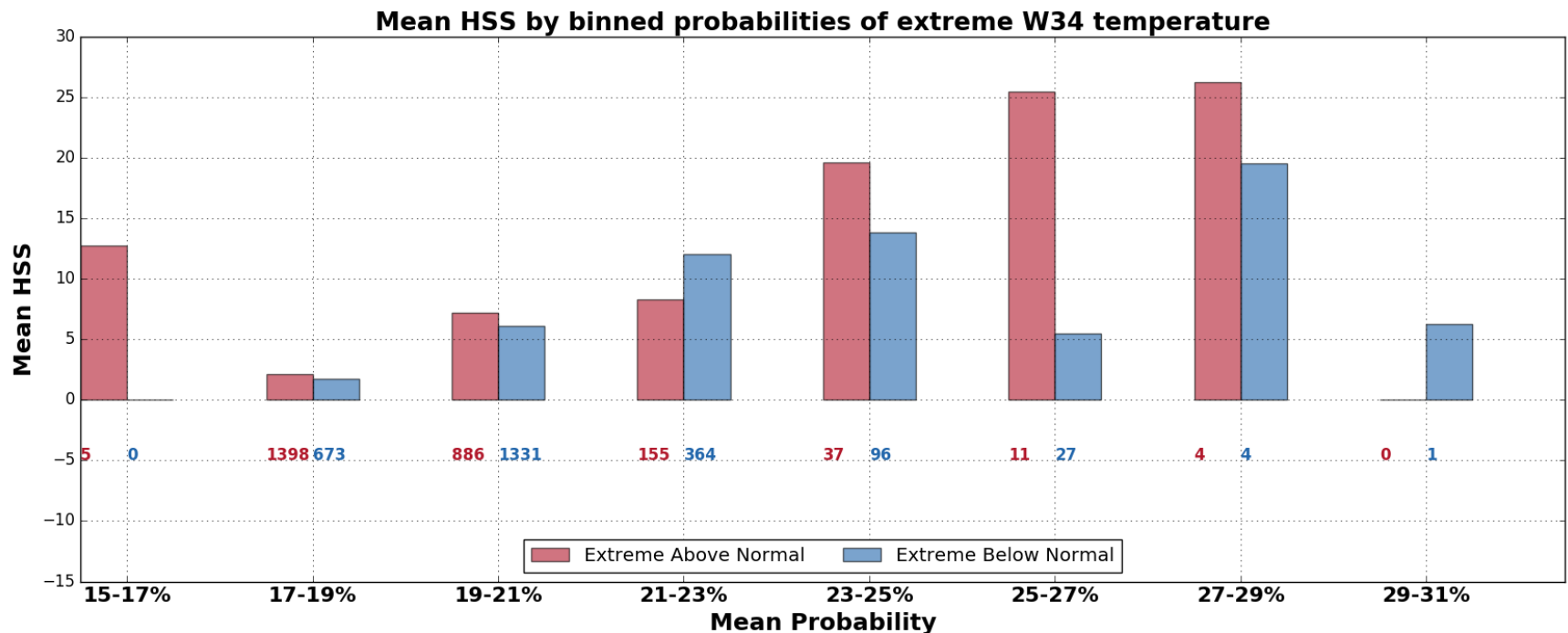
# Extreme **above** / **below** normal temperature

- **Calibration** improves Heidke Skill Score of **raw** extremes forecast
- **Greater probability** implies **greater skill**



# Finding forecasts with greater skill...

- Area-average calibrated forecast probabilities mostly range from **17-21 %**
- **Greater area-average probability produces greater skill**
  - Up to at least 25% probability for below normal extremes
  - Up to 29% for above normal extremes



# Summary

- Calibration at week 3-4 timescale essential to produce reliable probabilities for extremes
- Average skill is low, but there is intermittent skill
- Using a multi-model ensemble (MME) improves capacity to capture the extremes
- **Uncalibrated MME is overconfident for extremes**
- Higher probability forecasts represent periods of greater skill, or forecasts of opportunity
- Continued work to optimize identification of forecasts of opportunity