How Skillful are Collective El Niño-Southern Oscillation (ENSO) Predictions

Michelle L'Heureux (NOAA CPC), Rohan Shroff (Cornell U.), and Michael Tippett (Columbia U.)

Current ENSO Team: Emily Becker (CIMAS/Miami)*, Tom Di Liberto (NOAA CPO)*, Azhar Ehsan (IRI), Zeng-Zhen Hu, Nathaniel Johnson (NOAA GFDL)*, Arun Kumar, Michelle L'Heureux, Wanqiu Wang, Caihong Wen, Jieshun Zhu
* ENSO blog affiliates

47th Climate Diagnostics and Prediction Workshop
Logan, Utah
25-27 October 2022

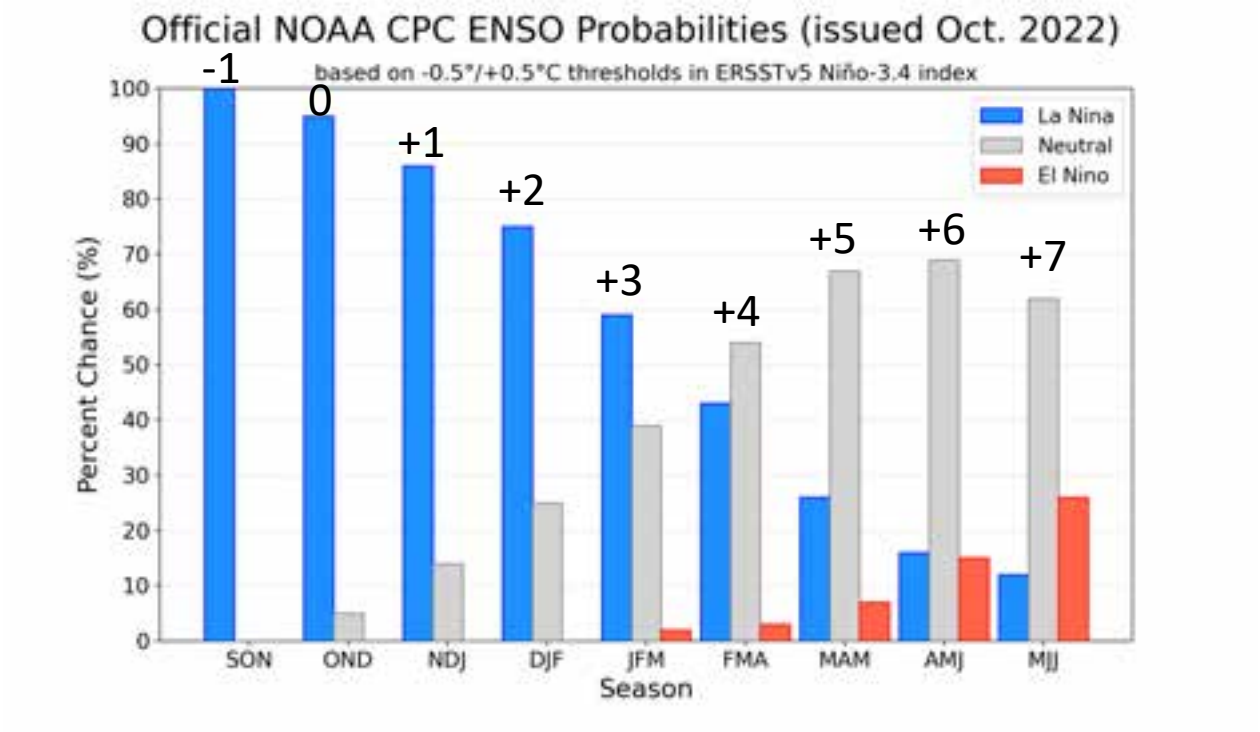
Background

- Team of 10-13 forecasters whose Oceanic Nino Index, ONI (seasonal average of Nino3.4 index) forecasts are averaged together to produce the Official, *Consensus* outlook.

- CPC started making these predictions in June 2015 so we could compute probabilities of a *strong* El Niño in 2015-16. Simply providing 3-category probabilistic forecasts for El Niño-Neutral-La Niña wasn't adequate.

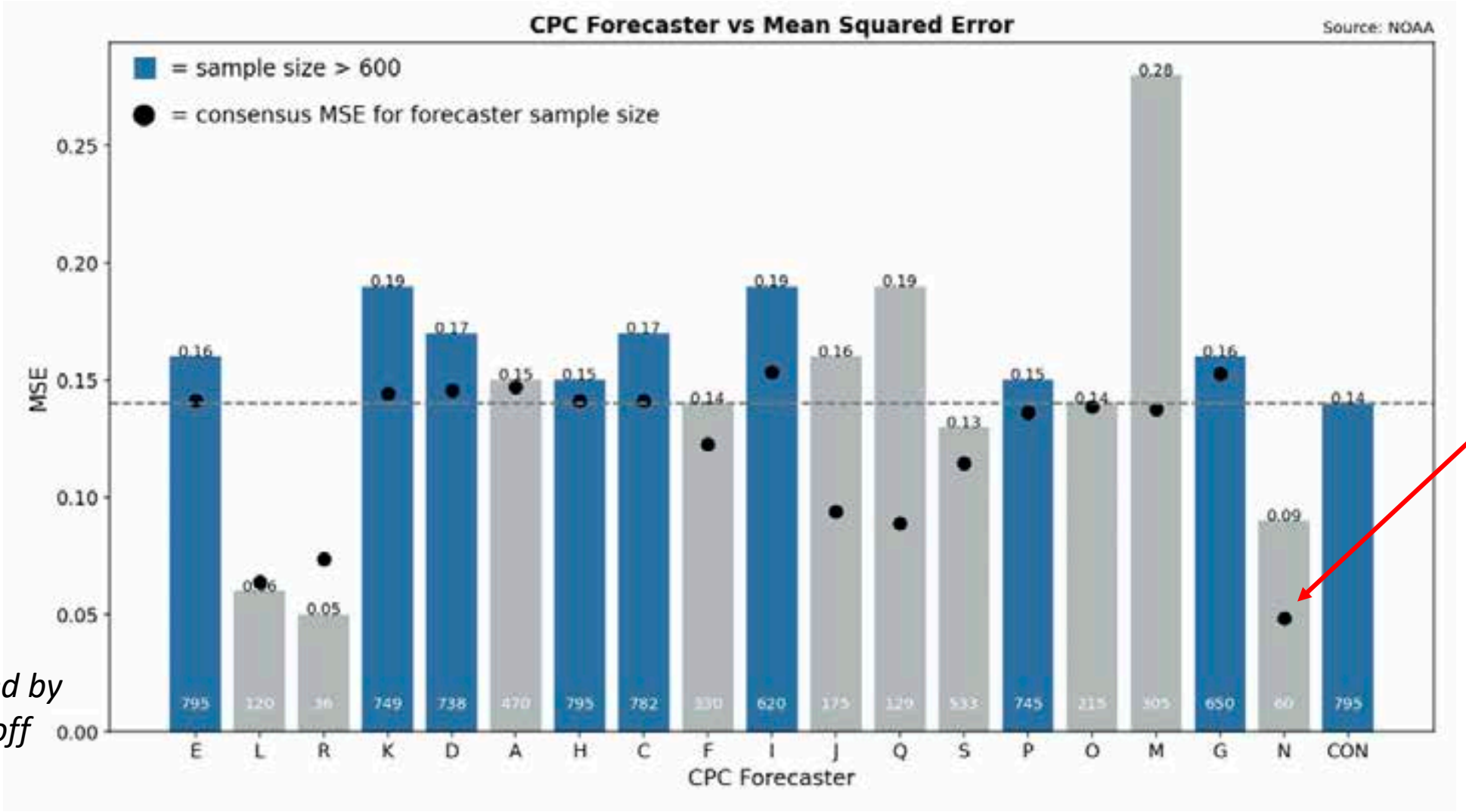
- 7 years of real-time forecasts (2015-2022)

- The team provides forecasts beginning at -2 lead time up to +7 lead time (publicly, we display -1 lead onwards).



Forecast issued on the **2nd Thursday** of every month.

What we Learned this Past Summer….



Dots is the Consensus error for times when forecaster made a prediction.

Dots lower than bars mean consensus error is smaller than individual.

Created by R. Shroff

- The Consensus offers smaller mean squared error (MSE) than the individual forecasters who contribute to it.
- *We're stronger together!*

Some Follow On Questions…

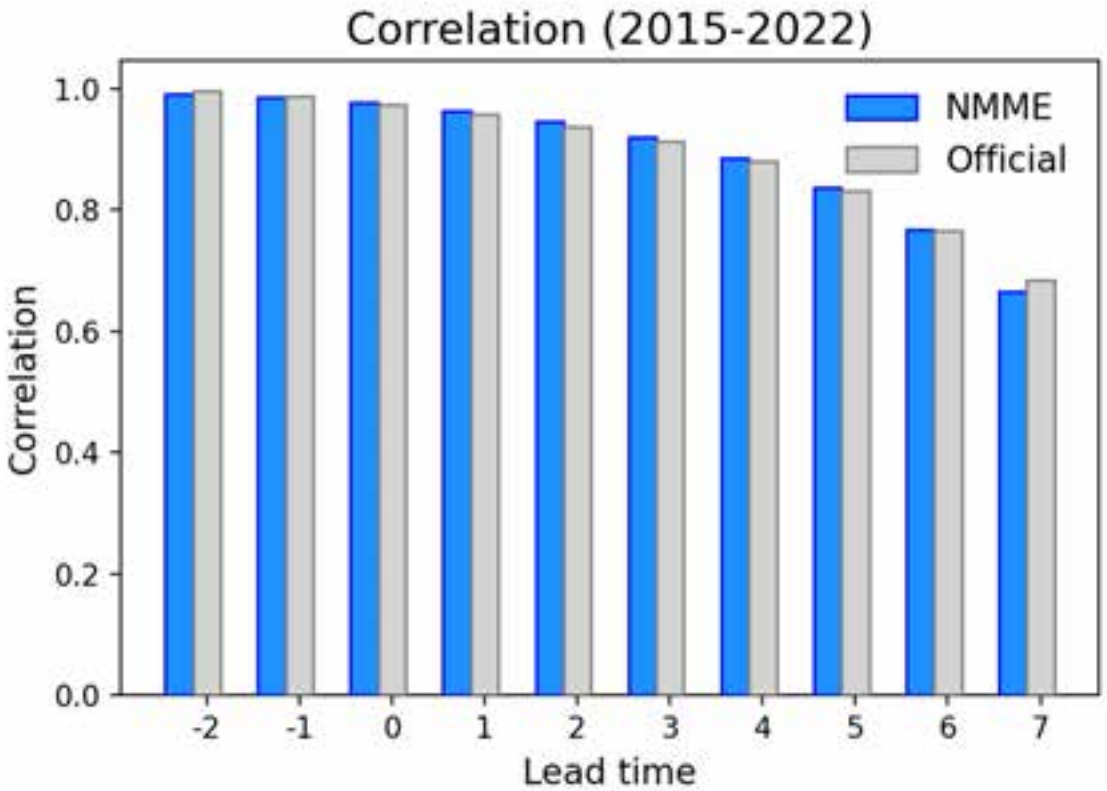
(1) How does the Consensus (or Official) average forecast compare to the North American Multi-Model Ensemble (NMME) average?

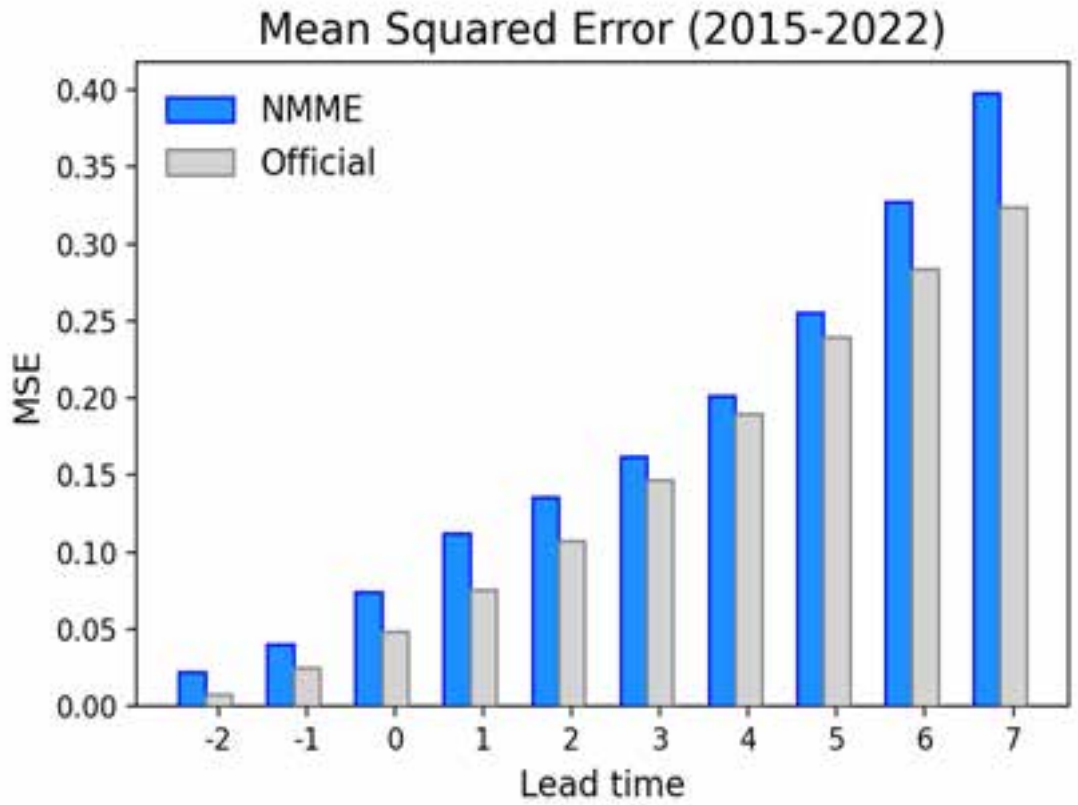
(2) If there are differences (spoiler: there are), what may be driving Consensus forecasts to be different from the NMME forecasts?

(3) Can we make the forecasts better?

(1) How does the Consensus (or Official) average forecast compare to the North American Multi-Model Ensemble (NMME) average?

-- North American Multi-Model Ensemble (NMME) is an average of 6 participating seasonal climate models
-- important point of comparison because we receive NMME right as we make the ENSO forecast.
-- expect the Consensus/Official forecast to look similar to the NMME…





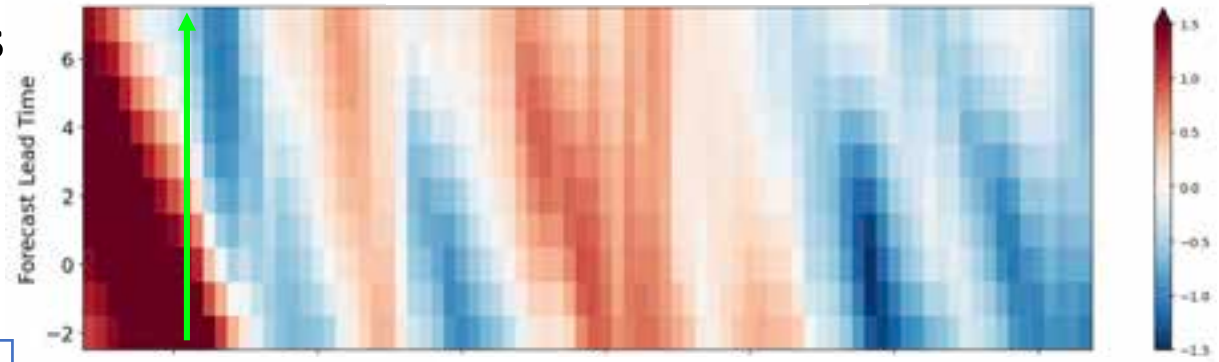
Why does the Consensus Forecast have smaller error (MSE)?

Consensus and NMME predictions are very similar.

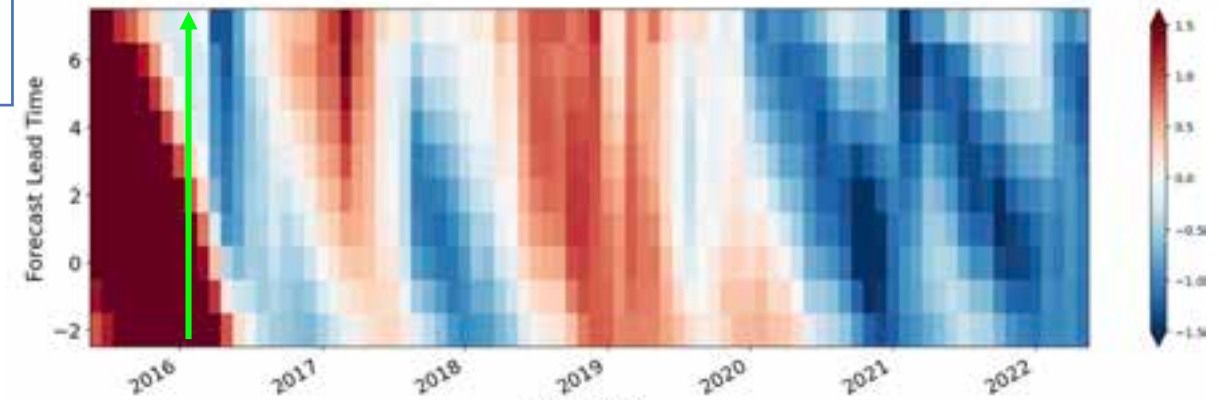
In this 7 year sample : 2 El Niño winters and 4 La Niña winters

Predictions made in early January 2016 out to +7 season lead

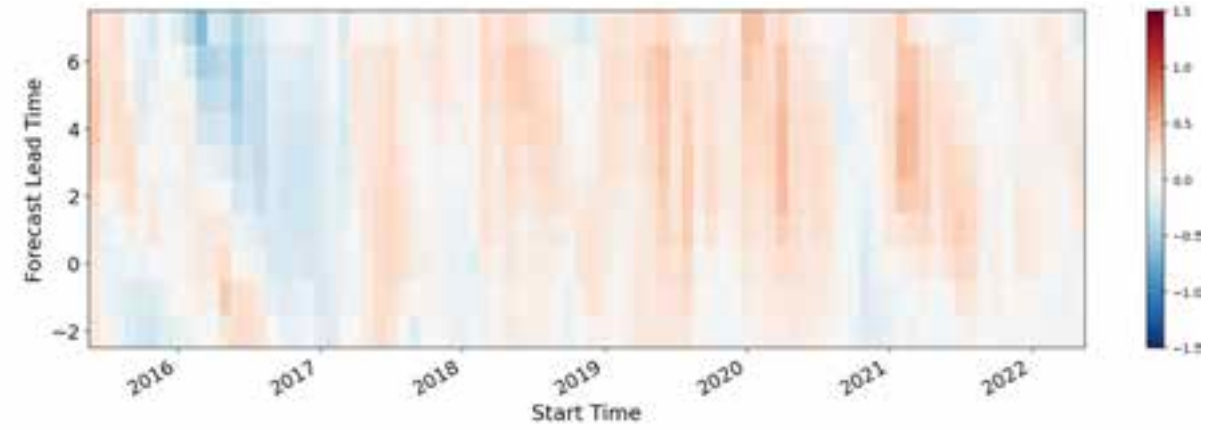
Consensus Forecast



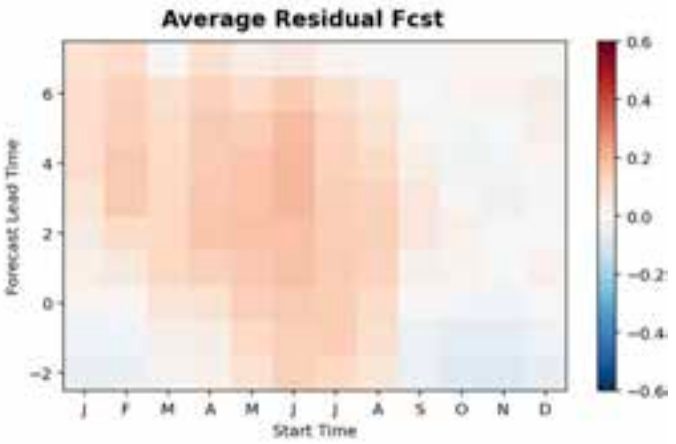
NMME Forecast



Residual Forecast (Consensus with linear NMME signal subtracted out)



Residual forecast is on the warmer side. Interesting.



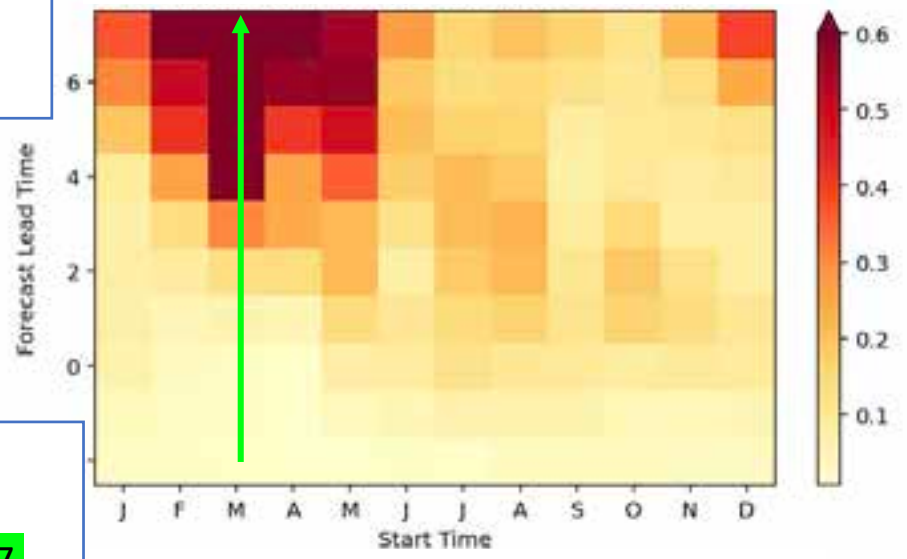
Q: What does the Consensus forecast look like when it is *uncorrelated* with NMME (r=0)?

(1) How does the Consensus average forecast compare to the North American Multi-Model Ensemble (NMME) average? Any seasonal differences?

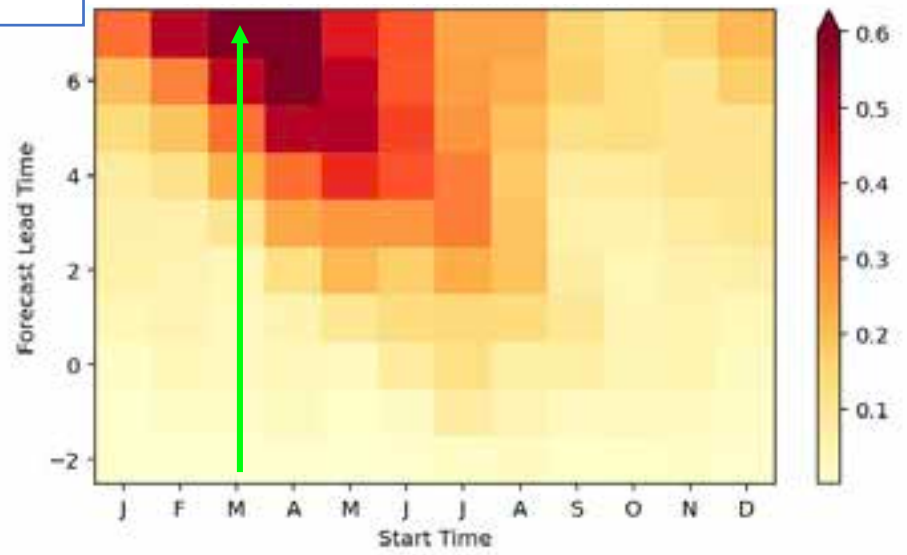
Bigger numbers = Larger Error

All predictions made in early March out to +7 season lead.

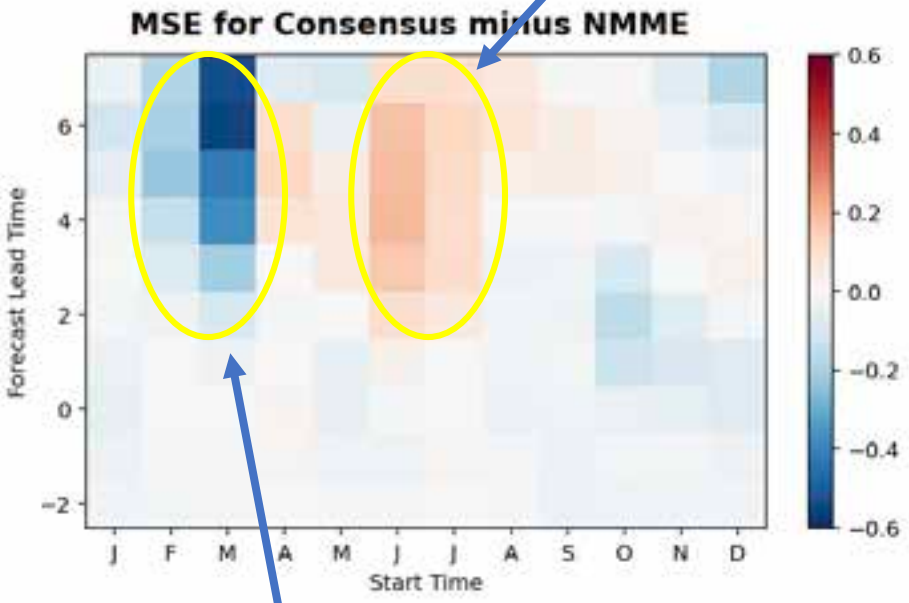
Mean Squared Error for NMME



Mean Squared Error for Consensus



Consensus error is larger (worse) than NMME for fcsts made in June/July



Consensus error is smaller (better) than NMME for forecasts made in Feb/Mar

(2) If there are differences, what may be driving Consensus forecasts to be different from the NMME forecasts?

Skill Scores Based on the Mean Square Error and Their Relationships to the Correlation Coefficient

ALLAN H. MURPHY

Department of Atmospheric Sciences, Oregon State University, Corvallis, Oregon

(Manuscript received 1 February 1988, in final form 11 April 1988)

SS is Mean Squared Error Skill Score (or MSESS).

$$\mathrm{SS}(f, \bar{x}, x) = r_{fx}^2 - [r_{fx} - (s_f/s_x)]^2 - [(\bar{f} - \bar{x})/s_x]^2. \quad (12)$$

correlation

variance ratio term

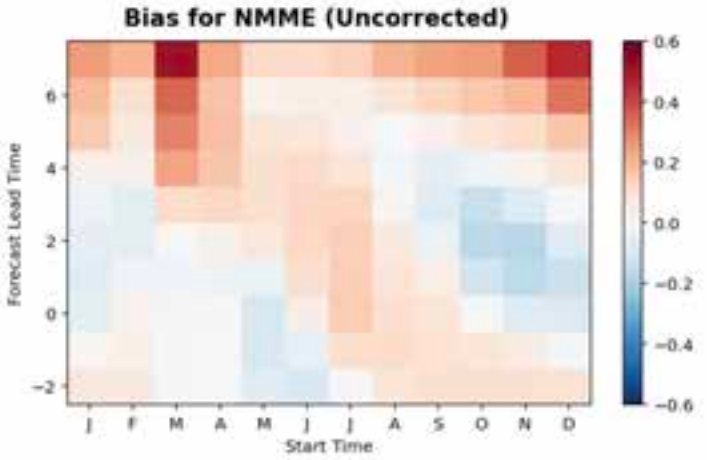
bias term

(M. Tippett reminds us that the exact same breakdown works for Brier Skill Score too)

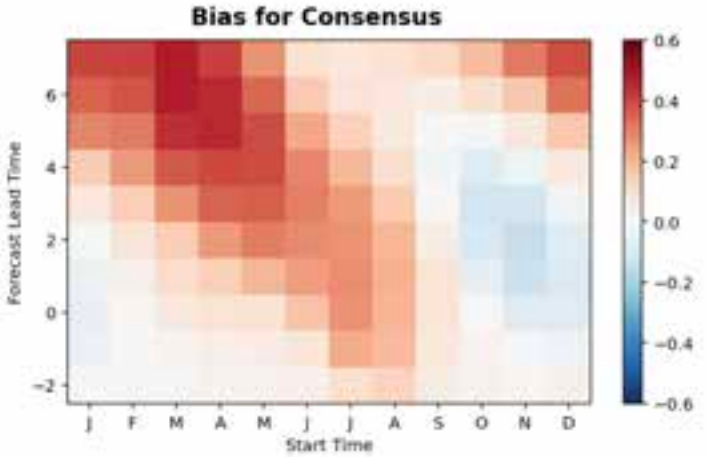
What does bias look like?

Red: Forecast mean is warmer than observed mean.

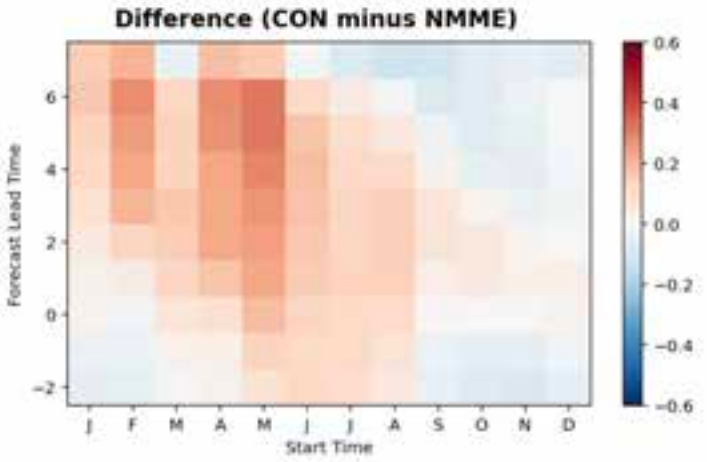
Blue: Forecast mean is colder than observed mean.



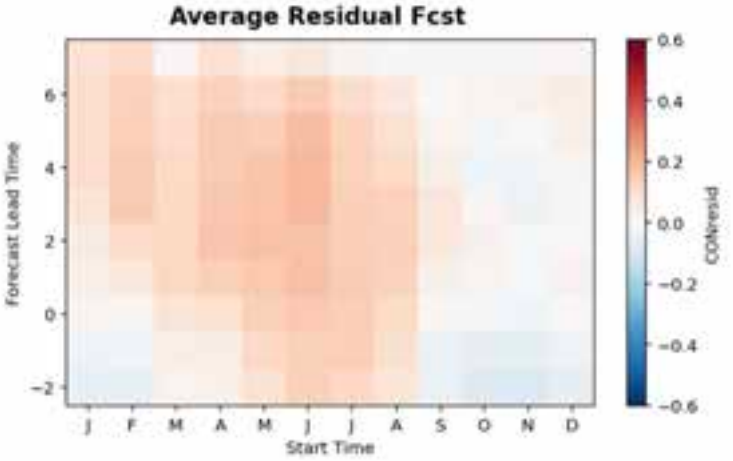
NMME too warm at the longest leads.



<u>Consensus has large biases!</u>
Too warm at longer leads and for Jan-Aug start times.



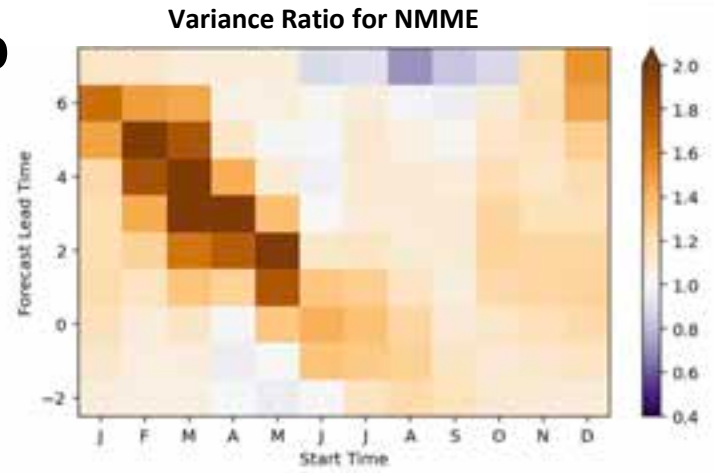
Recall, this was our plot for the Consensus residual (uncorrelated with NMME). We were finding bias.



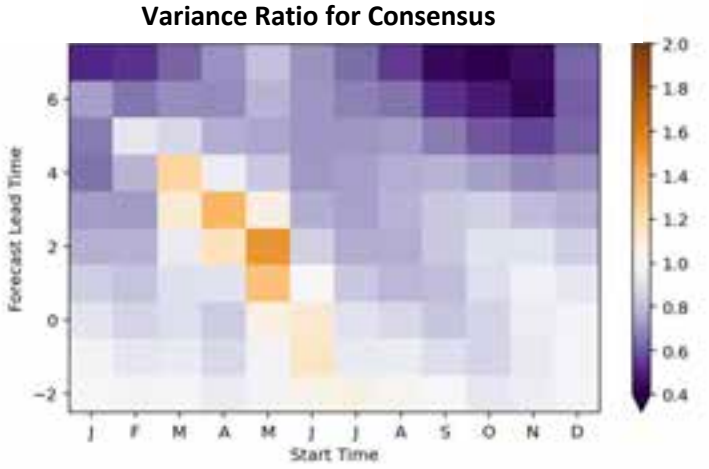
What does the variance ratio look like?

Orange: Forecast variance is more than observed variance

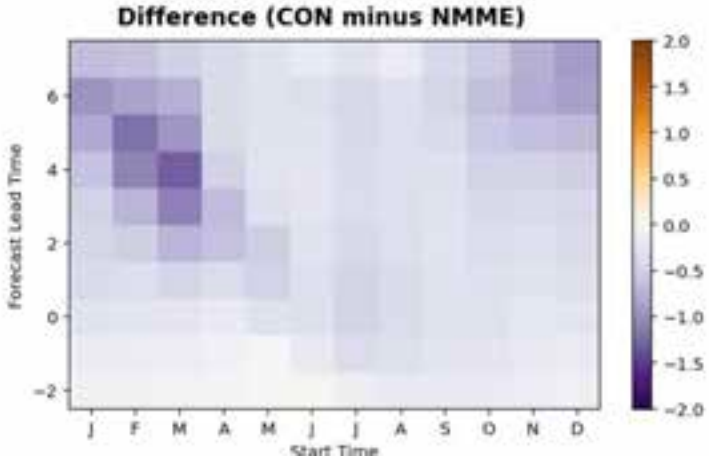
Purple: Forecast variance is less than observed variance



NMME fcst variance is too large!



Consensus has generally smaller forecast variance than observations.



(2) If there are differences, what may be driving Consensus forecasts to be different from the NMME forecasts?

Skill Scores Based on the Mean Square Error and Their Relationships to the Correlation Coefficient

ALLAN H. MURPHY

Department of Atmospheric Sciences, Oregon State University, Corvallis, Oregon

(Manuscript received 1 February 1988, in final form 11 April 1988)

SS is Mean Squared Error Skill Score (or MSESS).

$$\mathrm{SS}(f, \bar{x}, x) = r_{fx}^2 - [r_{fx} - (s_f/s_x)]^2 - [(\bar{f} - \bar{x})/s_x]^2. \quad (12)$$

To get a large skill score → Want a large correlation

Want to make this variance ratio term as small as possible.

To make this zero:
$r = (S_f / S_x)$

Want to make this bias term as small as possible.

To make this zero:
$\bar{f} = \bar{x}$

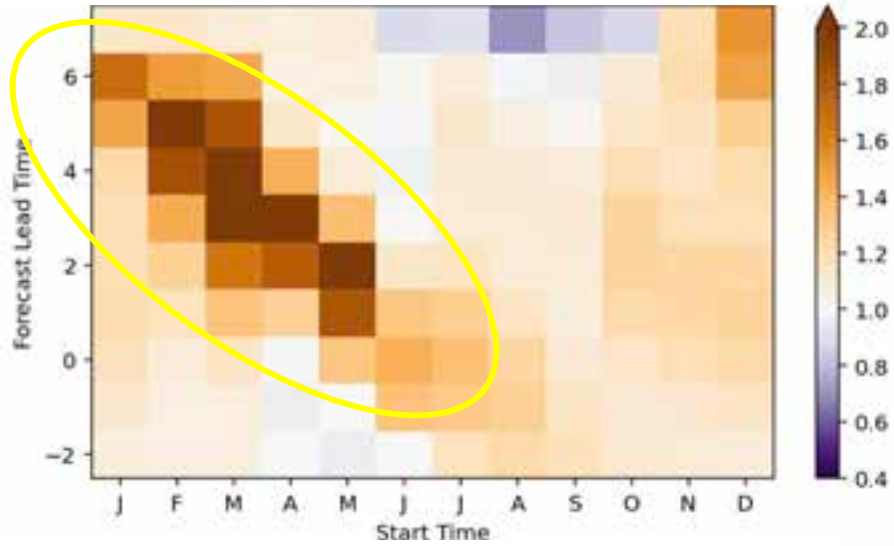
What <u>should</u> the variance ratio look like?

NMME forecast variance too large given the skill.

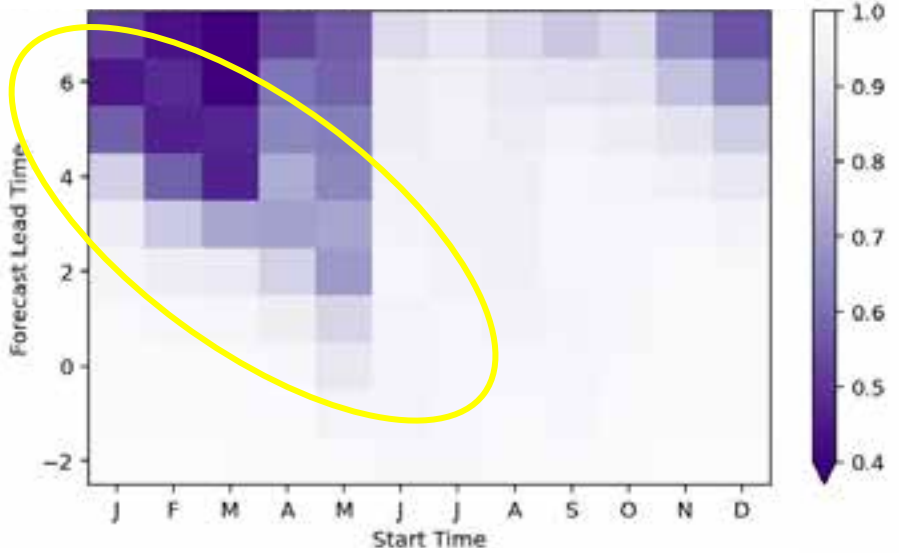
Consensus (human) intuitively understands forecast variance should be smaller.

But we are too cautious in the 2nd half of the year.

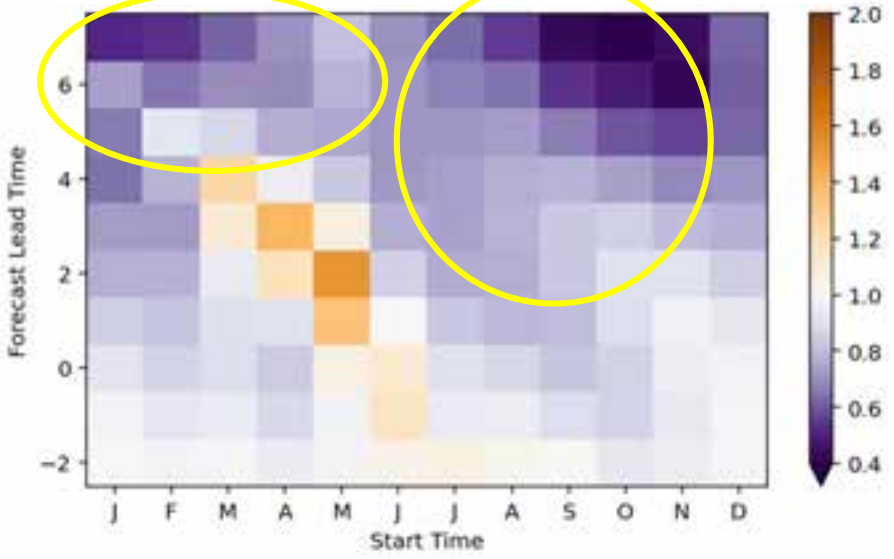
Variance Ratio for NMME



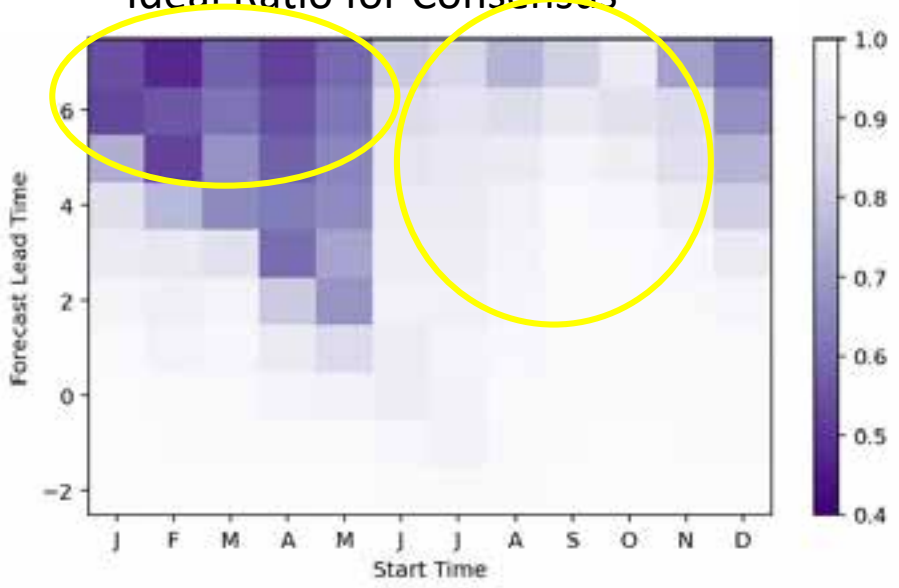
Ideal Ratio for NMME



Variance Ratio for Consensus

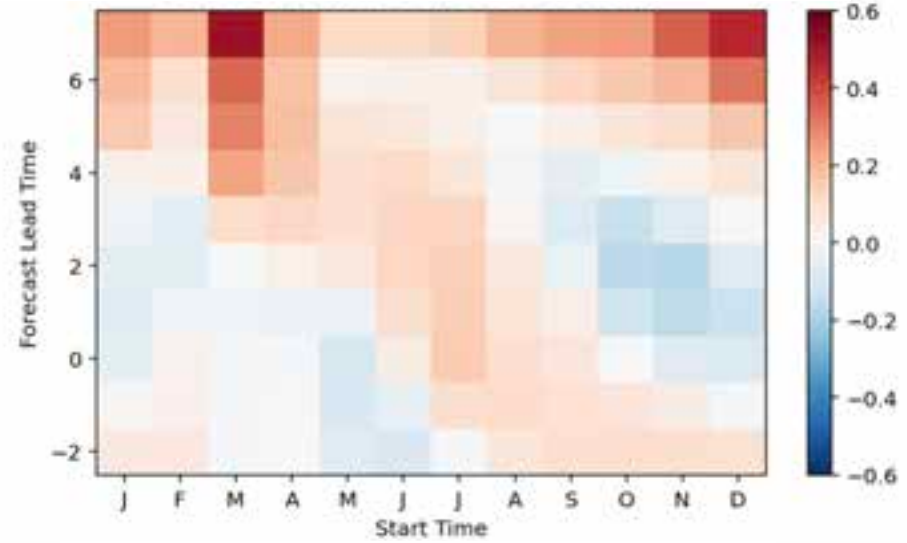


Ideal Ratio for Consensus

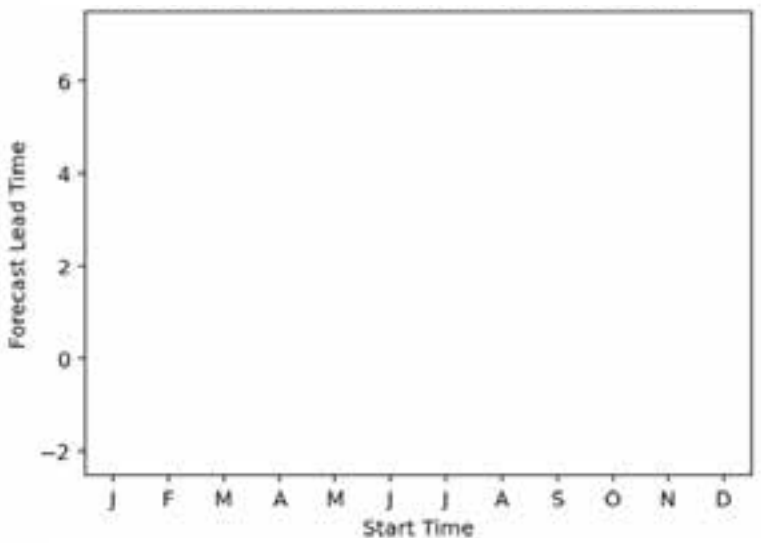


What <u>should</u> the bias look like?

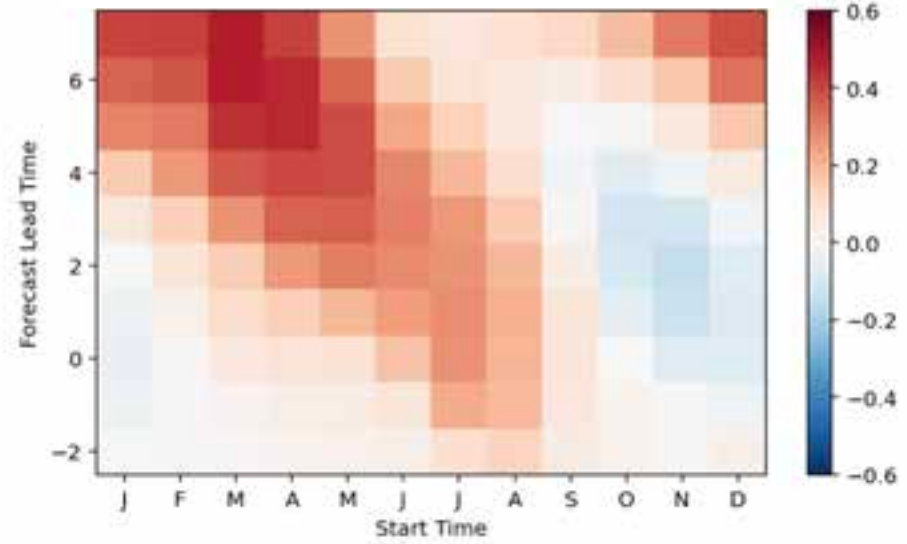
Bias for NMME



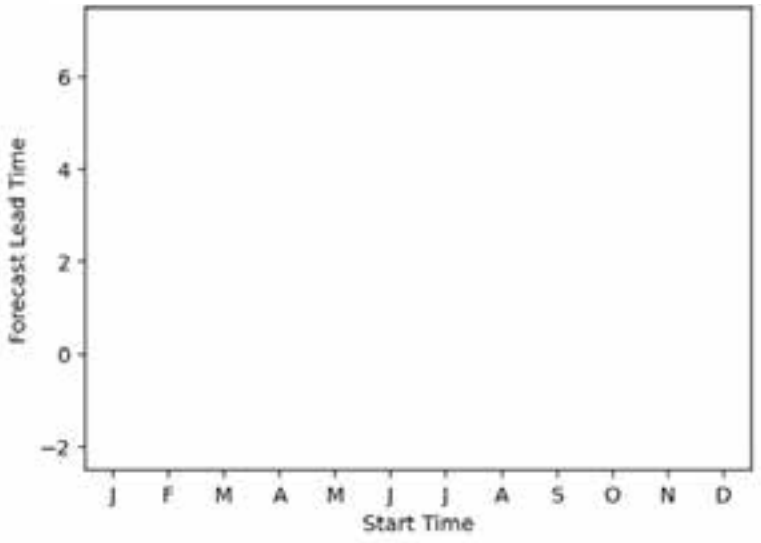
Ideal Bias for NMME



Bias for Consensus



Ideal Bias for Consensus

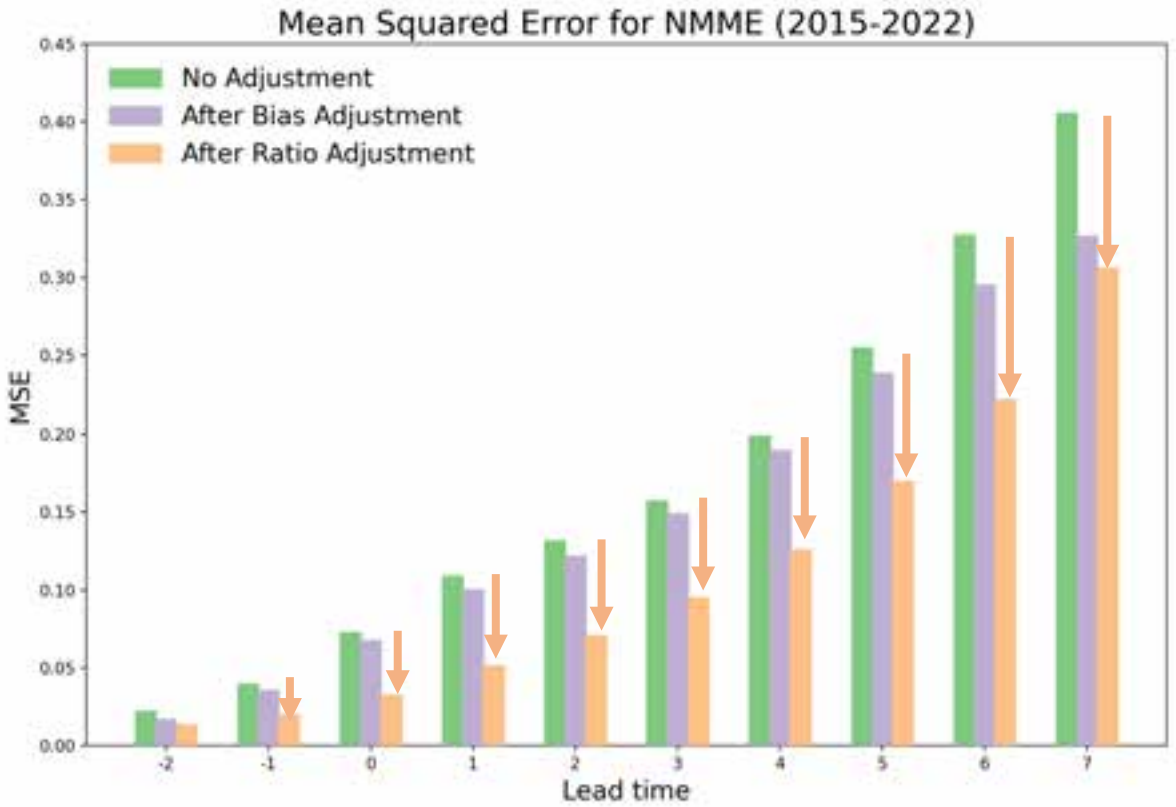


Role of Sampling?
Too warm in this period with 4 La Niñas.

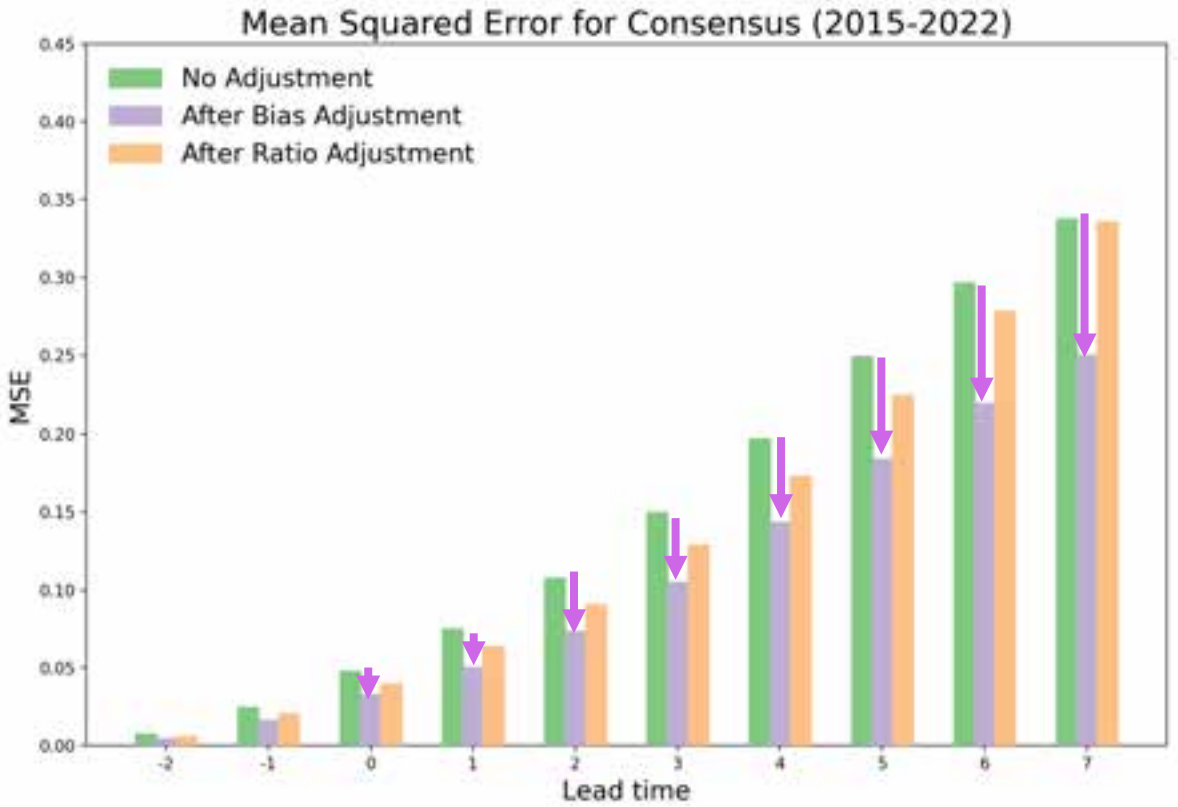
(3) Can we make the forecasts better?

-- Yes! We can adjust the forecast by multiplying by the regression coefficient and subtracting the bias.

-- regression coefficient is equal to correlation if forecast variance = observed variance, otherwise you need to scale by the *variance ratio*.



NMME error decreases more after *the variance* is corrected.



Consensus error decreases more after *the bias* is corrected.

Note: Not cross validated.

Conclusions:

(1) How does the Consensus (or Official) average forecast compare to the North American Multi-Model Ensemble (NMME) average? Any seasonal differences?

-- Correlation is nearly the same. Consensus has smaller errors than NMME (especially for predictions made in Jan-Mar).

(2) If there are differences, what may be driving Consensus forecasts to be different from the NMME forecasts?

-- NMME errors are mostly attributable to forecast variance errors (too large) and Consensus errors are mostly attributed to forecast bias (too warm) errors.
-- *likely sampling plays a role here with only 7 years of data.*

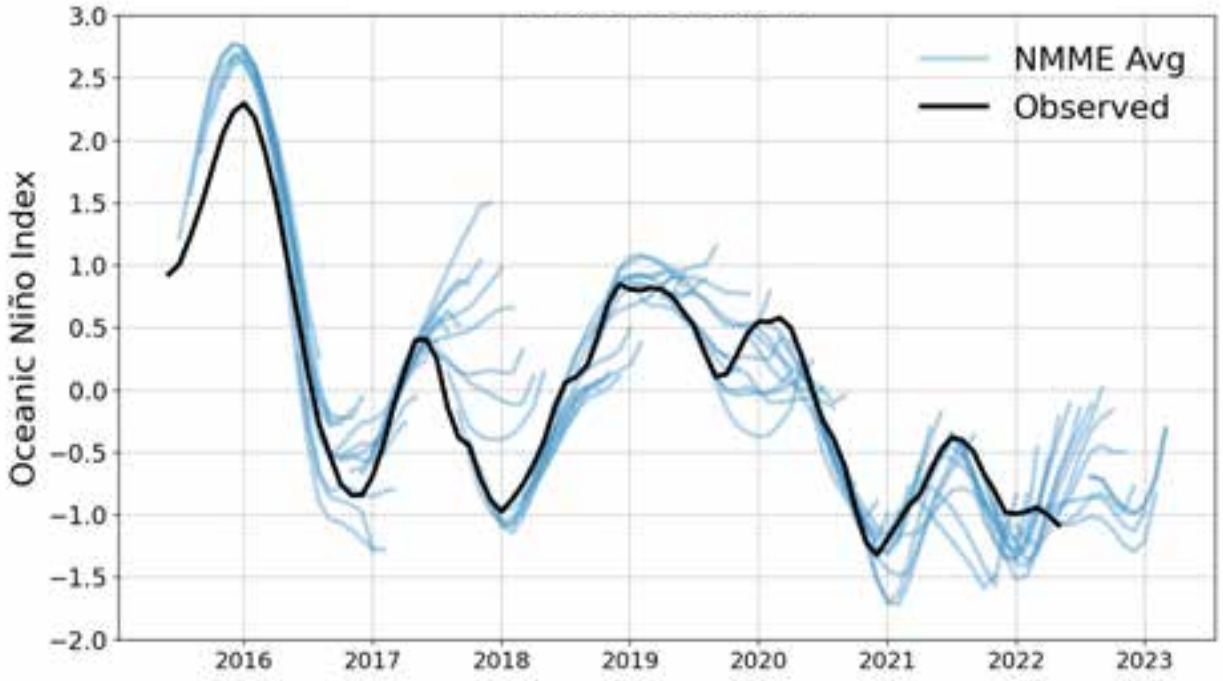
(3) Can we make the forecasts better?

-- yes, regression-based methods are simple and effectively correct for variance and bias errors.
-- consensus (humans) need to check their biases
-- model forecasts need to reduce their amplitudes (especially in the early spring)

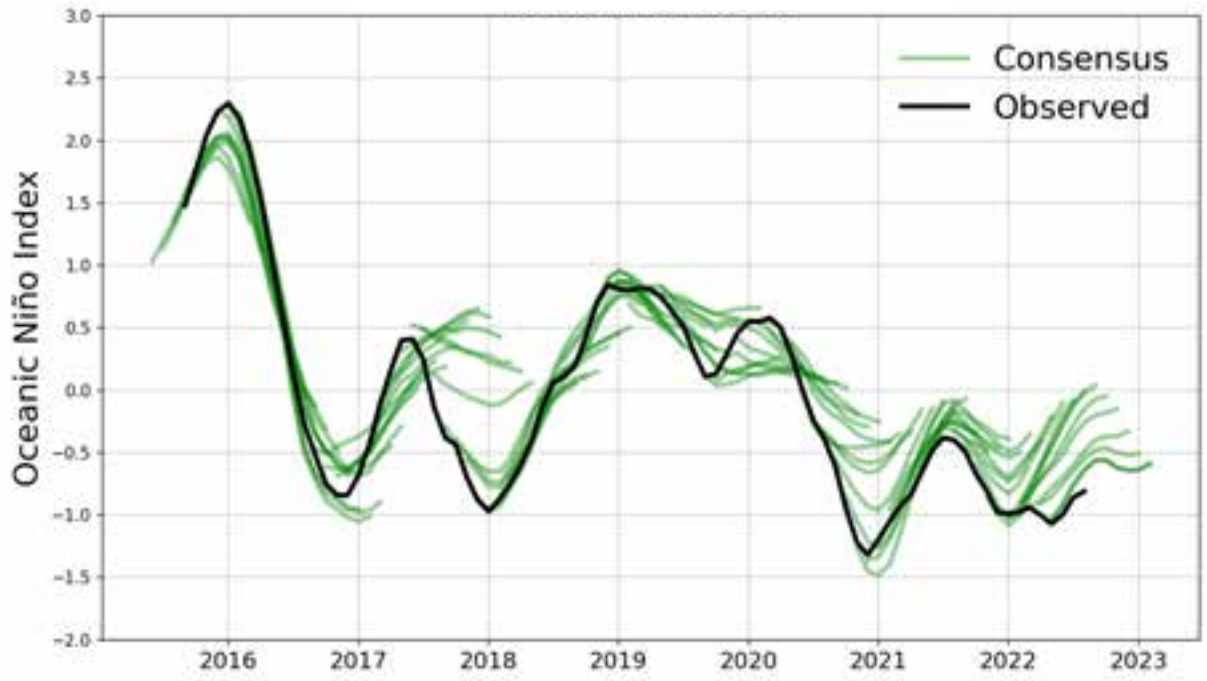
Extra Slides

(1) How does the Consensus average forecast compare to the North American Multi-Model Ensemble (NMME) average?

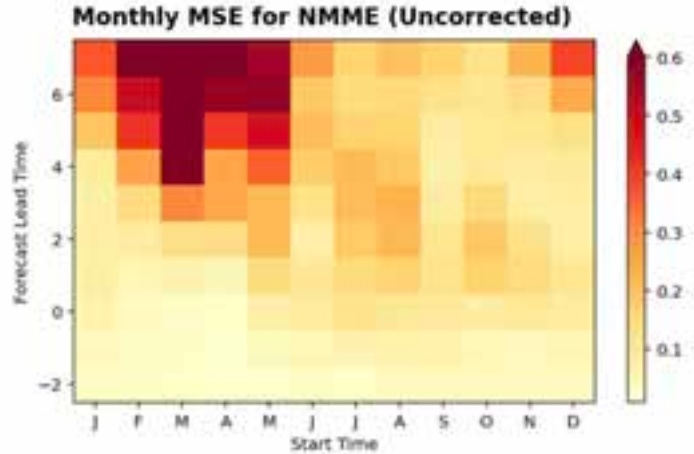
NMME Average Forecast of ENSO

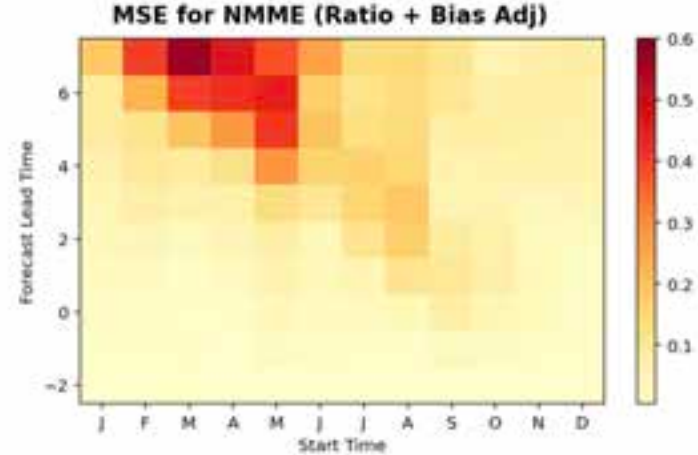


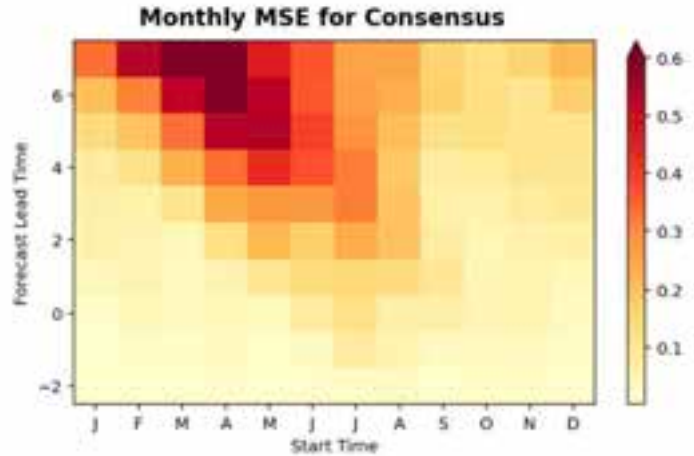
Consensus Forecast of ENSO

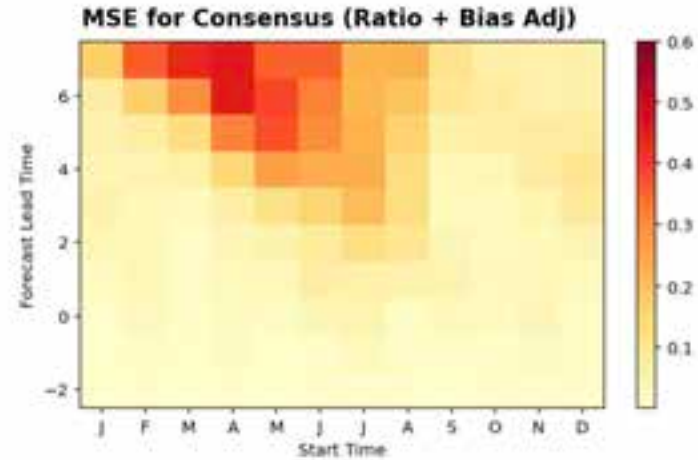


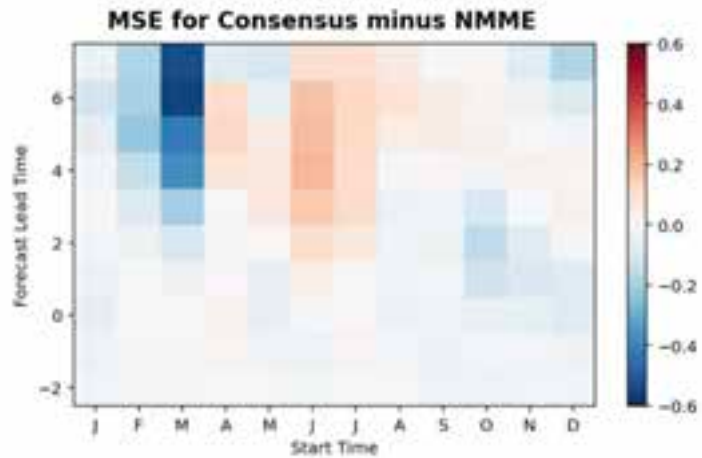
(3) Can we make the forecasts better?

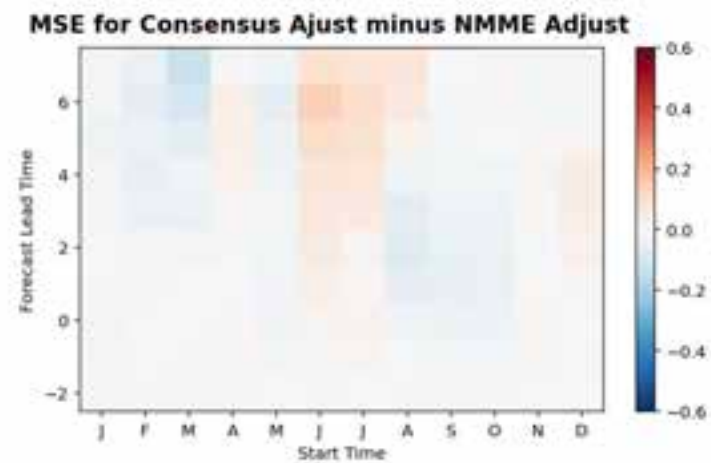


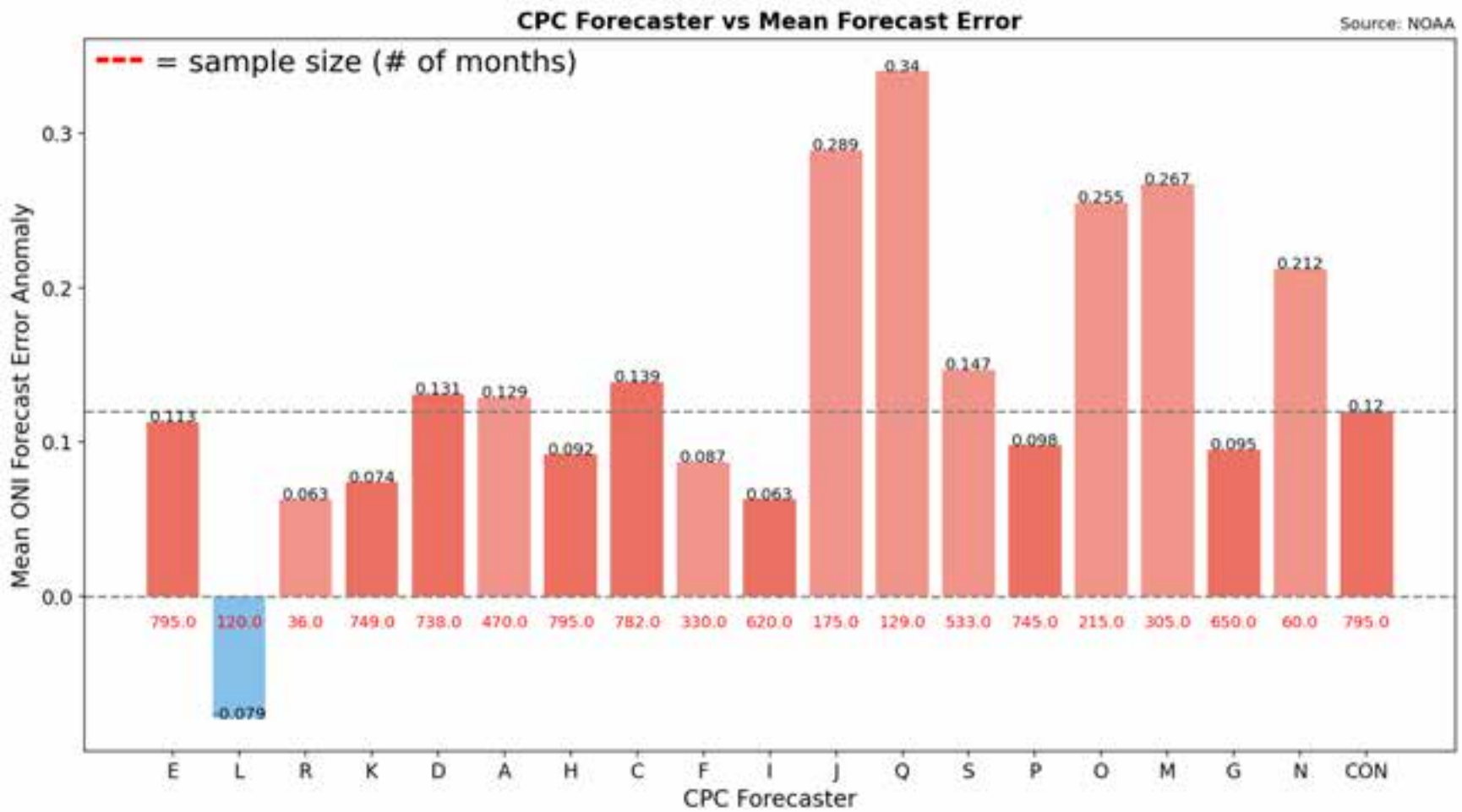











CPC Forecaster vs Mean Forecast Error
Source: NOAA
--- = sample size (# of months)
Mean ONI Forecast Error Anomaly
0.3
0.2
0.1
0.0
0.113
-0.079
0.063
0.074
0.131
0.129
0.092
0.139
0.087
0.063
0.289
0.34
0.147
0.098
0.255
0.267
0.095
0.212
0.12
795.0
120.0
36.0
749.0
738.0
470.0
795.0
782.0
330.0
620.0
175.0
129.0
533.0
745.0
215.0
305.0
650.0
60.0
795.0
E
L
R
K
D
A
H
C
F
I
J
Q
S
P
O
M
G
N
CON
CPC Forecaster