A Linear Inverse Model for Improved Week 3-4 CPC Operational Outlooks

John R. Albers[1]

Matthew Newman[1]

Yuan-Ming Cheng[1,2]

Maria Gehne[1,2]

[1]NOAA - Physical Sciences Laboratory

[2]Cooperative Institute for Research in the Environmental Sciences
University of Colorado Boulder

August 15, 2023

Topics for today:

- **Evaluate forecast skill of updated CPC-PSL linear inverse model**

  → LIM "v2.0" is about to be transferred to CPC servers (LIM v1.0 has been running as an operational S2S forecast guidance tool at CPC since late 2020)

  → Give particular attention to the LIM's forecast skill over California

- **Use research LIM to diagnose elevated summer-fall California forecast skill**

  → Discuss LIM-based 'expected skill' tool for identifying 'forecasts of opportunity' in real-time

  → Evaluate the role of soil moisture and SST in producing the LIM's elevated forecast skill

- **Future developments of the LIM for California needs, with new CPC-DWR support**

Models and verification data:

Operational models:

- Linear inverse model (LIM)

  12 (bimonthly) LIMs trained on 1958-2016 JRA-55 reanalysis data (detrended using a “fair-sliding” 20-yr retrospective mean)

    - Forecasts from LIMs are blended across adjacent months to yield seamless year-round deterministic and probabilistic predictions
- ECMWF IFS real-time forecasts (CY 43R1- 47R3) (S2S prediction database, Vitart et al 2017)
    - Mean bias correction applied using ‘on-the-fly’ 20-year hindcasts

Research models:

- 2 LIMs trained on:
    - AMJ 1959-2021 JRA-55 reanalysis data (15-fold cross-validated training/verification)
    - JAS 1959-2021 JRA-55 reanalysis data (15-fold cross-validated training/verification)

Verification data for all Week 3-4 forecasts:

- JRA-55 reanalysis 2m temperature
- 2 forecasts per week (year-round) for operational LIMs and IFS
- 7 forecasts per week for cross-validated hindcasts

Skill metrics:

- Anomaly correlation and anomaly pattern correlation

- Two category Heidke skill score (HSS) *(Peng/Kumar/Halpert/Barnston WAF 2012)*

  - HSS computed as both a time series (averaged over all United States grid points) and spatial map (each grid point averaged over all time)

$$\text{Heidke skill score (HSS)} = \frac{(c - e)}{(t - e)}$$

- $c \Longrightarrow$ number of grid points with hits
- $t \Longrightarrow$ total number of grid points in forecast
- $e \Longrightarrow$ number of grid points expected to be correct by chance (t/2)

<u>Interpretation:</u>

- HSS = 0 $\Longrightarrow$ Forecast skill equal to random chance
- HSS = 1 $\Longrightarrow$ Perfect forecast
- HSS < 0 $\Longrightarrow$ Forecast skill worse than random chance

Baseline assumption:

S2S prediction involves the evolution of climate anomalies (aggregates of weather, not individual weather events), where chaotic nonlinearities may largely be approximated as unpredictable noise

$$\frac{dx}{dt} = Lx + \xi$$

<u>"Slow" timescale processes</u>

- Explicitly model ("macroscale") processes that are potentially predictable on subseasonal timescales

<u>"Fast" timescales processes (LIM noise forcing)</u>

- Rapidly decorrelating (<1 week) daily timescale synoptic variability
- "Parameterization" of nonlinear processes that are unlikely to be predictable on subseasonal timescales
- Observationally constrained (function of lag-zero covariance statistics)

What is a linear inverse model (LIM)?

- Empirical model, where the forecast operator (L) is constructed from covariances of 7-day running mean anomalies of observational data (here Japanese Reanalysis JRA-55)

$$\frac{dx}{dt} = Lx + \xi$$

Forecasted variables:

$$x = \begin{bmatrix} p \\ \Phi \\ H \\ \psi_T \\ \psi_{UTLS} \\ SST \\ T_{2m} \\ S_w \end{bmatrix}$$

- Mean sea-level pressure (20˚-90˚N)
- Geopotential height (500 hPa, 20˚-90˚N)
- Tropical heating (-14˚S-14˚N)
- Tropospheric stream function (750 hPa, 20˚-90˚N)
- Upper troposphere-lower stratosphere geopotential height (100 hPa, 30˚-90˚N)
- Tropical sea surface temperature (-14˚S-14˚N)
- 2m temperature (North America-land only)
- “Root zone” soil wetness (first two layers - North America-land only)

State independent white noise

$$\xi$$

- Observational constrained
- Function of L, x

Baseline forecast skill for United States (CONUS and Alaska)

(skill measured by HSS)

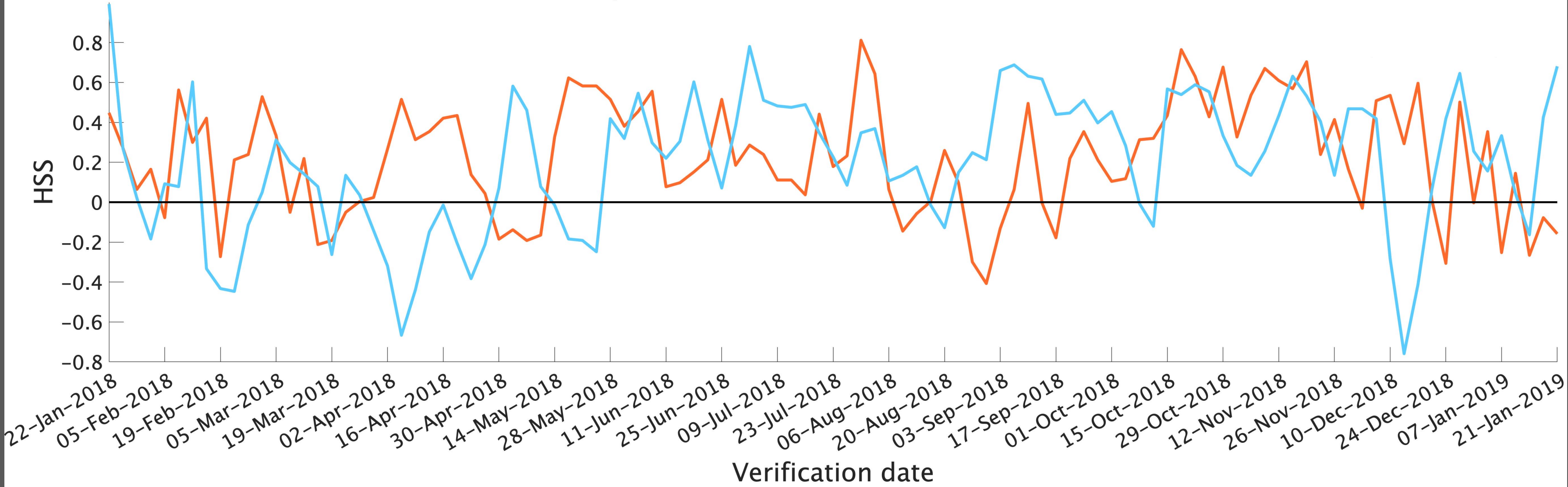
	2017	2018	2019	2020	2021	5-year average: (2017-2021)
CPC-PSL LIM	0.23	0.2	0.24	0.33	0.27	0.25
ECMWF IFS	0.23	0.22	0.31	0.32	0.27	0.27

Week 3-4 2m temperature HSS

(CONUS + Alaska)

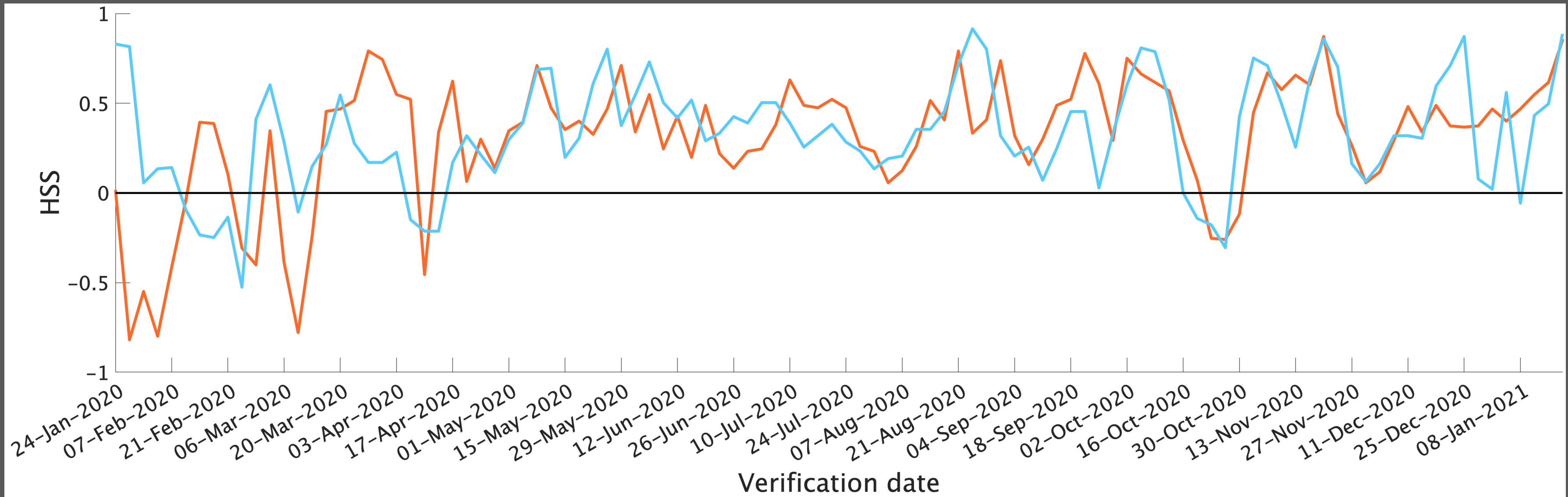
2018

- ECMWF IFS (mean HSS: 0.224)
- CPC LIM (mean HSS: 0.2)



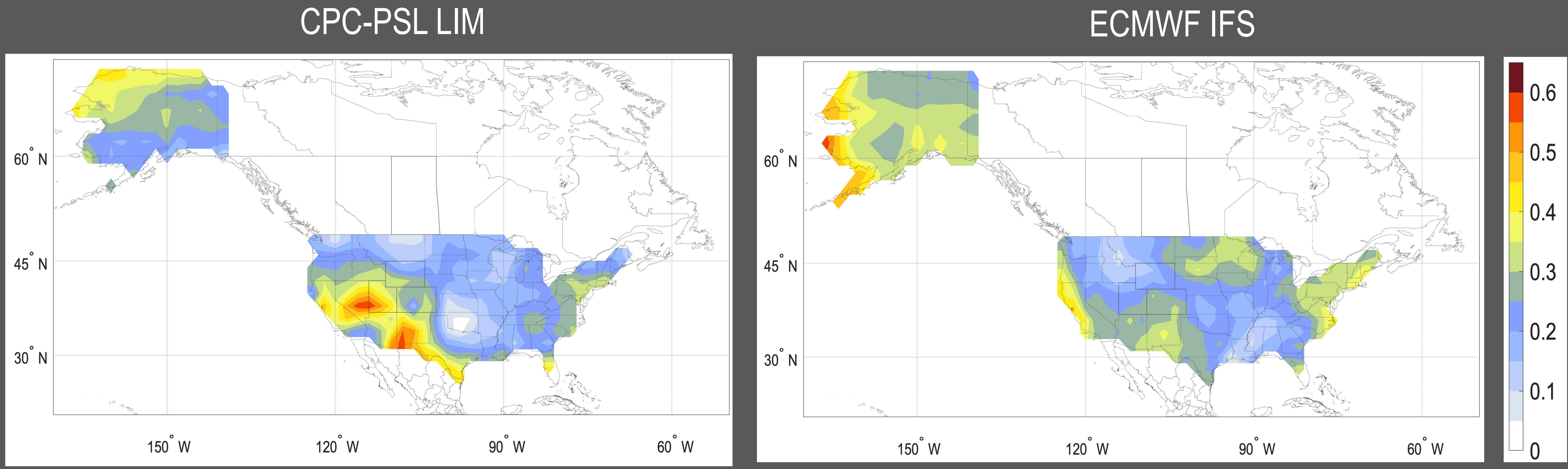
2020

- ECMWF IFS (mean HSS: 0.315)
- CPC LIM (mean HSS: 0.334)



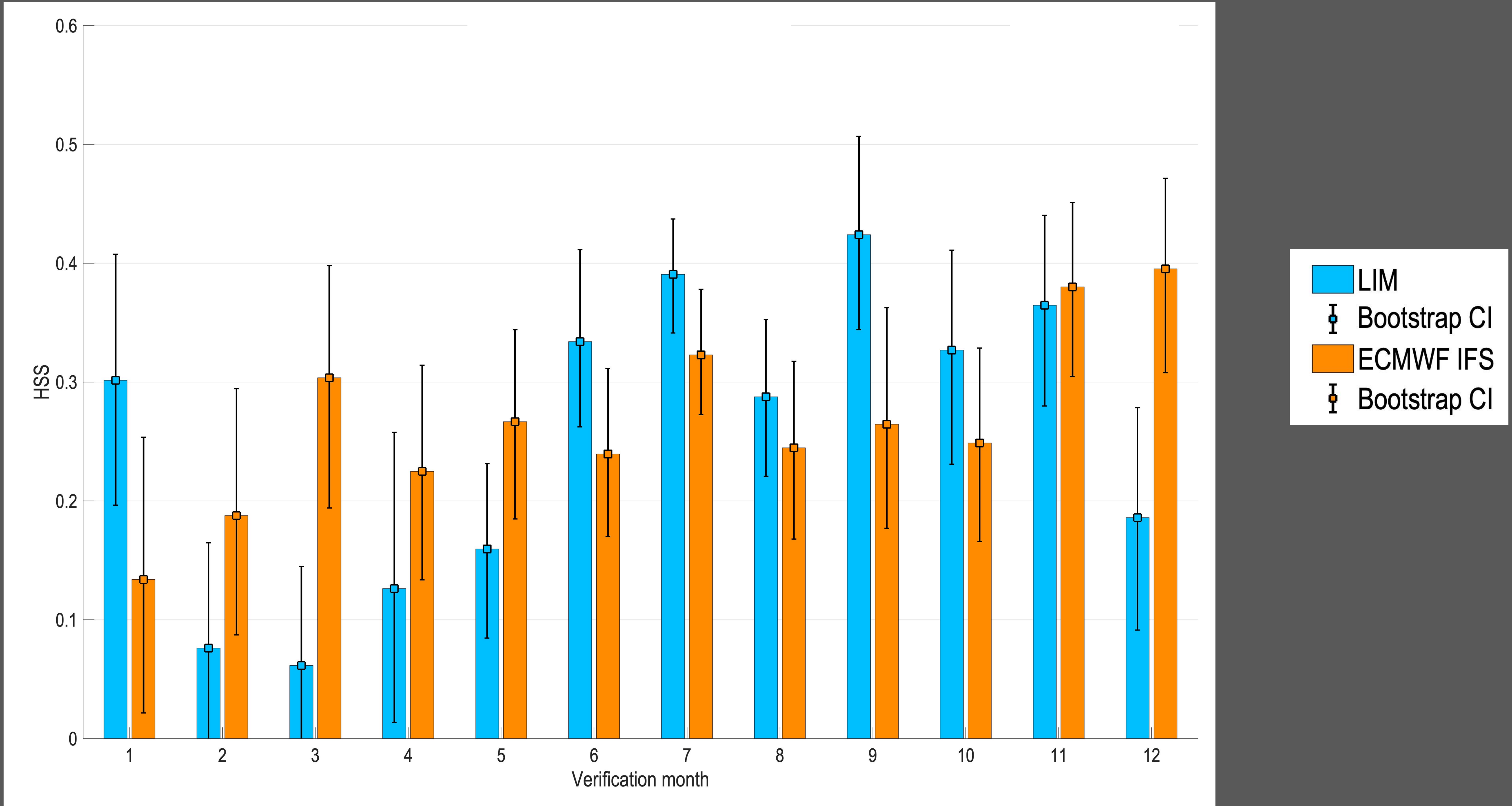
Week 3-4 2m temperature HSS

(Year-round 2017-2021)



Week 3-4 2m temperature HSS

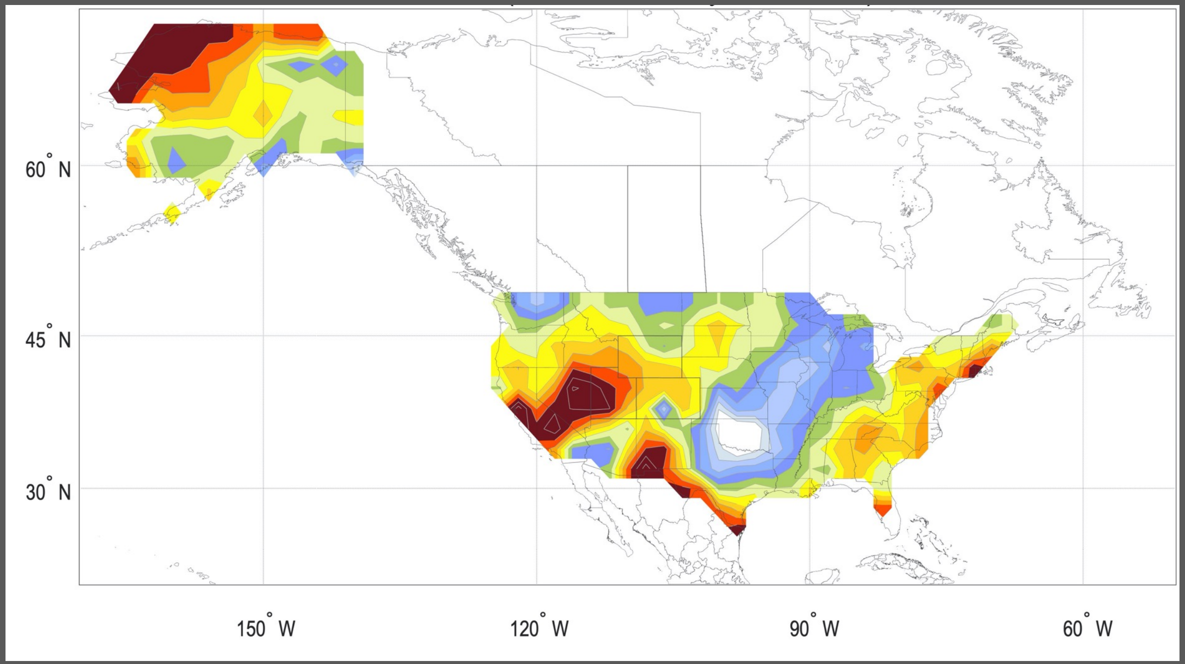
(2017-2021 by verification month)



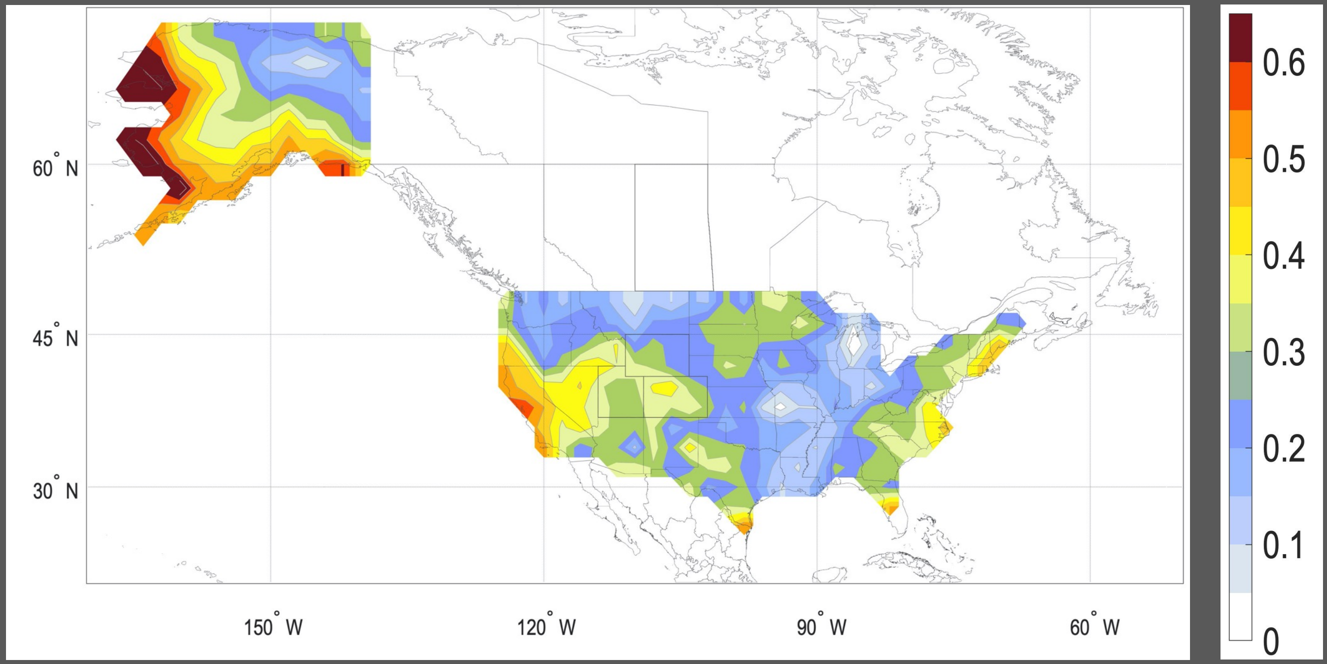
Week 3-4 2m temperature HSS

(late June – early November - 2017-2021)

CPC-PSL LIM



ECMWF IFS



Summary of baseline skill assessment:

- In aggregate, across all seasons and regions (CONUS+Alaska), the Weeks 3-4 2m temperature skill of ECMWF IFS and CPC-PSL LIM is roughly equivalent
- IFS slightly more skillful in spring, LIM more skillful in summer/fall
- LIM appears to be more skillful over the western US (including CA)

Questions:

1. How important is soil moisture to summer-fall forecast skill?
   - Experiment 1: Rerun research LIM hindcasts with soil moisture initial conditions set to zero

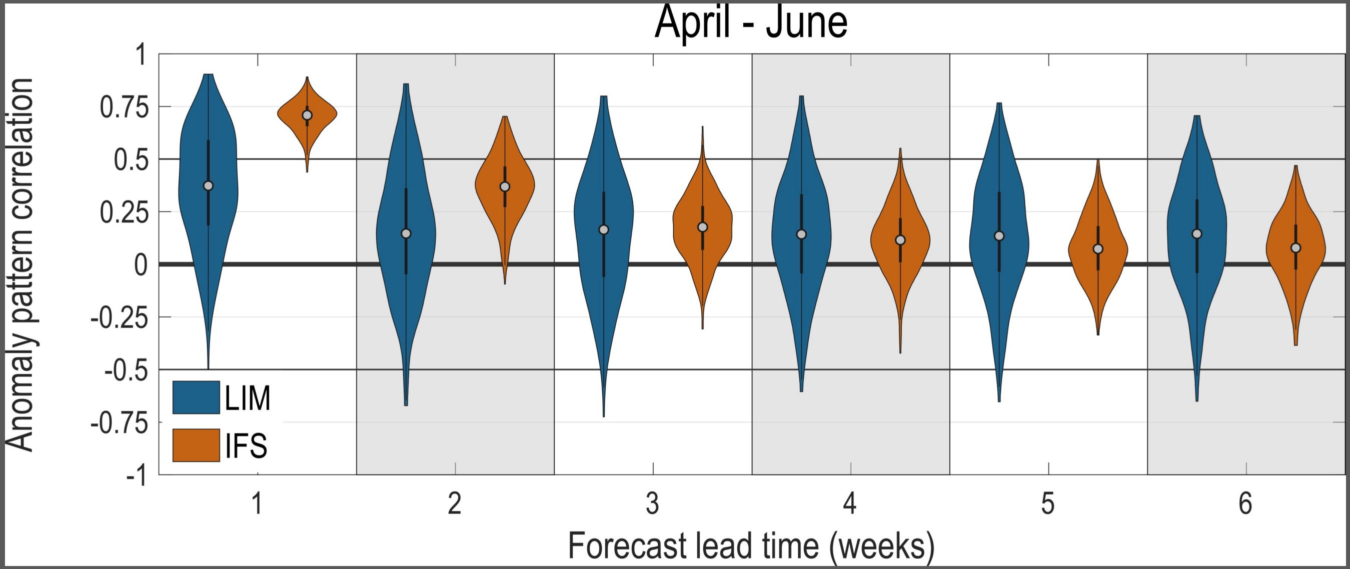
2. Are there other processes that modulate the soil moisture-temperature relationship?
   - Experiment 2: Rerun research LIM hindcasts with dynamical modes related to co-evolving tropical SST-North American soil moisture anomalies are suppressed

3. Can we identify higher skill 'forecasts of opportunity' related to soil moisture?
   - *Use LIM-based "expected skill" to predict when forecasts of opportunity will occur*

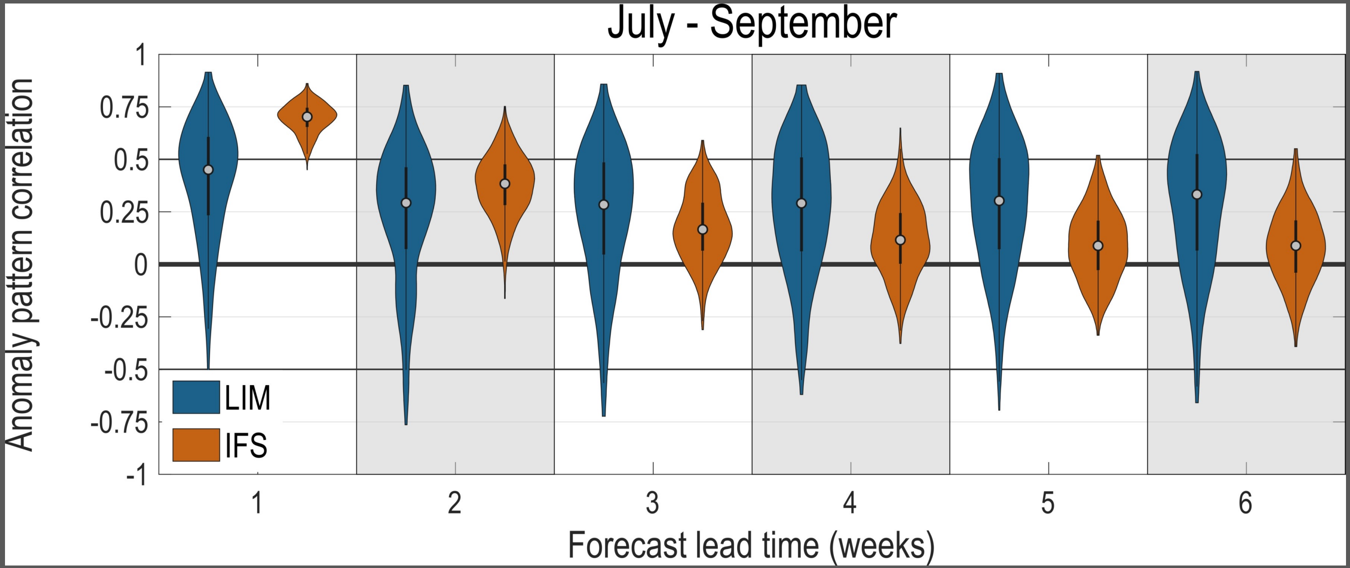
** expected skill is a function of LIM forecast signal-to-noise ratio*

North American 2m temperature hindcast skill (1997-2016)

(skill measured by pattern correlation)



- IFS more skillful for Weeks 1 and 2
- IFS and LIM roughly equivalent skill for Weeks 3 to 6



- IFS more skillful for Weeks 1 and 2
- LIM more skillful for Weeks 3 to 6

Note: LIM skill here is cross-validated hindcasts not outside of training period

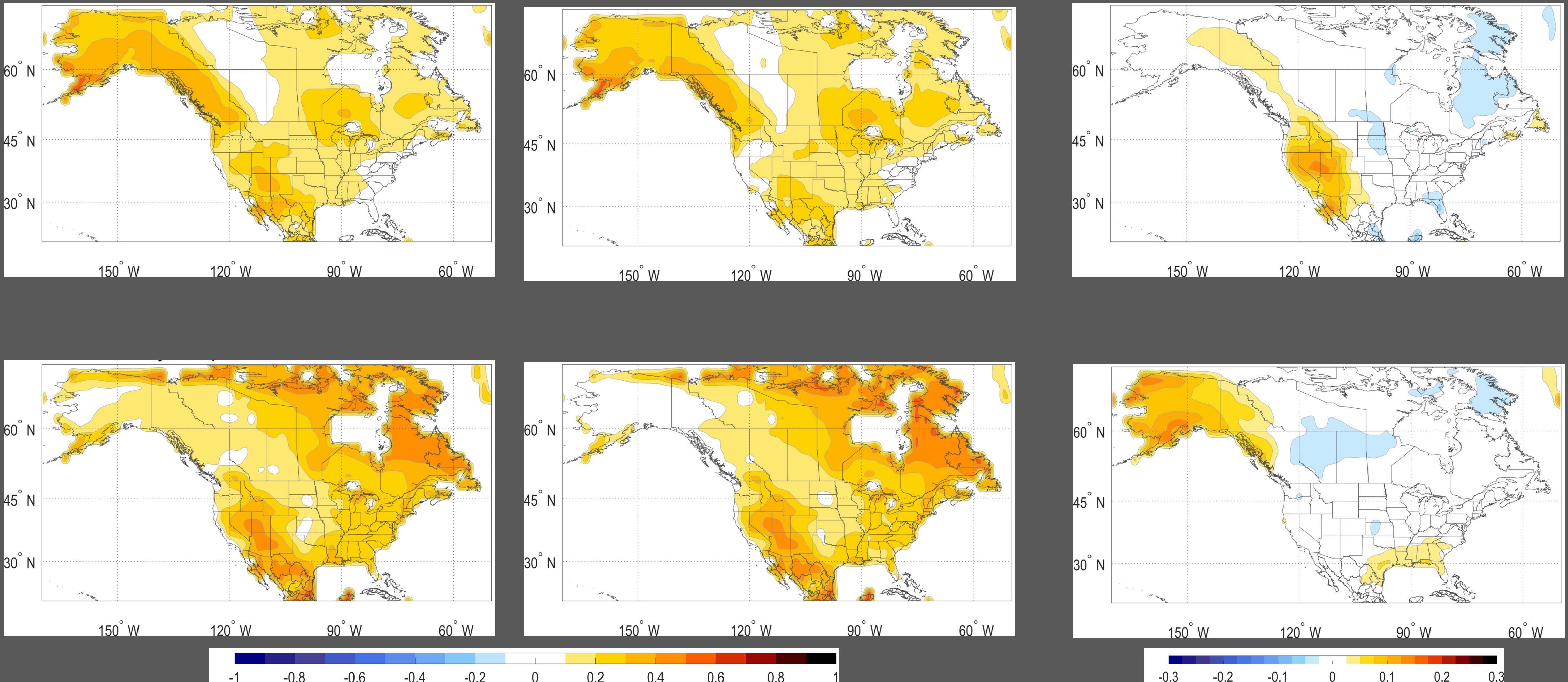
2m temperature forecast skill with/without soil moisture in initial condition

(April-June and July-September 1959-2021)

With soil moisture | Without soil moisture | Difference

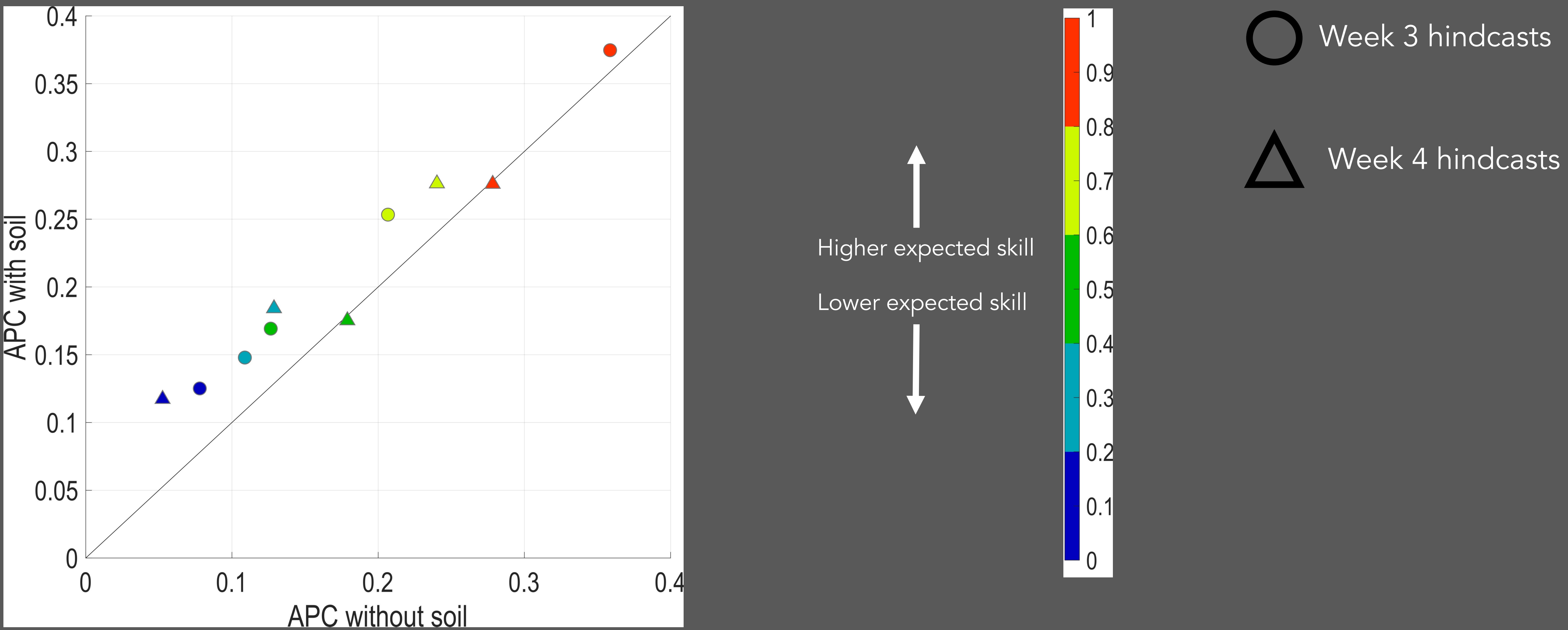
AMJ

JAS



Does soil moisture improve skill of forecasts of opportunity?

Western US hindcast skill with and without soil initial condition

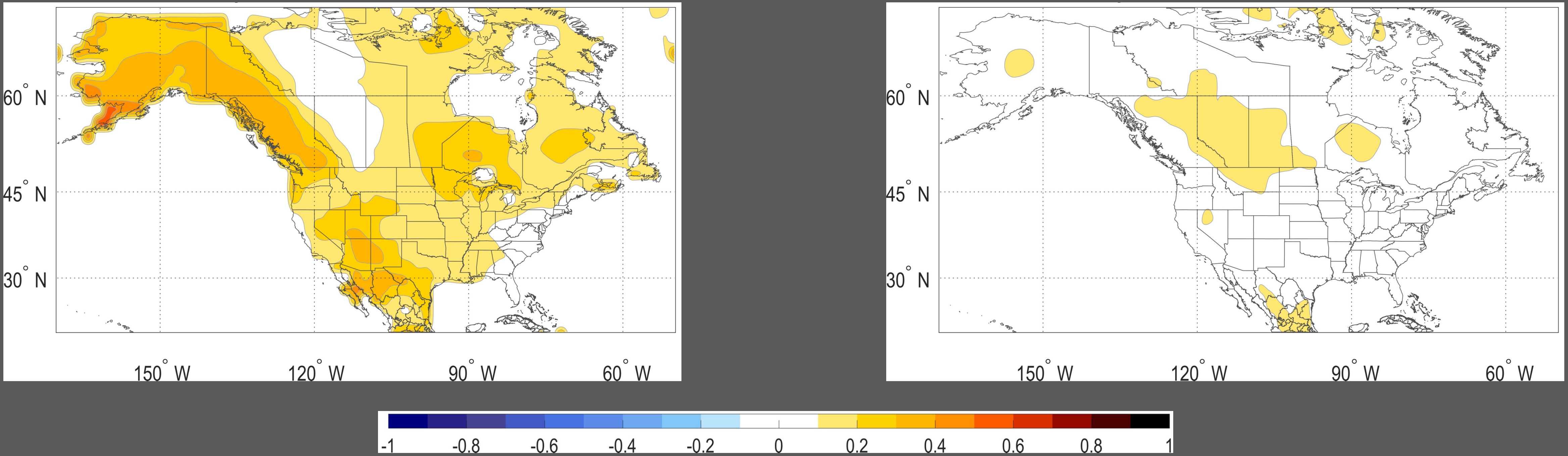


⟹ Soil moisture primarily improves skill of low-to-moderate skill forecasts

Dynamically filtered Weeks 3-4 2m temperature hindcasts
(AMJ 1959-2021)

Unfiltered hindcasts

Hindcasts filtered to exclude dynamical modes related to joint tropical SST-soil moisture anomalies



⟹ Virtually all week 3-4 2m temperature skill is associated with anomalies that involve the co-evolution of tropical SSTs and soil moisture

How does a LIM identify forecasts of opportunity?

LIM forecast signal

$$\frac{dx}{dt} = Lx + \xi$$

LIM noise forcing
(forecast uncertainty)

'*Expected skill*' of a perfect model infinite-member ensemble mean forecast

$$\rho_\infty(t;\tau) = \frac{S^2(t;\tau)}{\left(\left[S^2(t;\tau)+1\right]S^2(t;\tau)\right)^{1/2}}$$

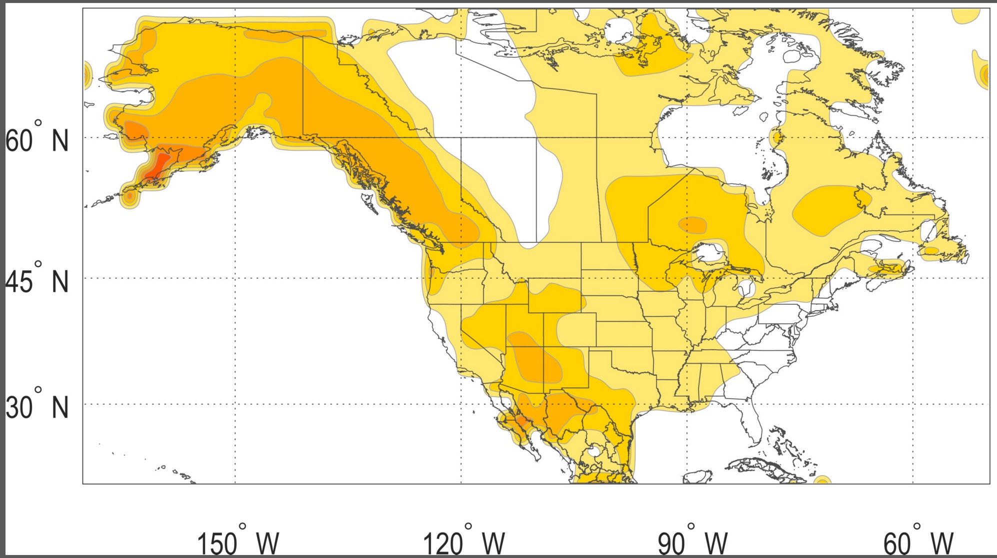
- $S^2 \longrightarrow$ forecast signal-to-noise ratio (based on the LIM in our case)
- $t \longrightarrow$ forecast initial time
- $\tau \longrightarrow$ forecast lead

- Calculated at time of forecast (it is a *forecast of forecast skill*)
- Forecast lead dependent

(Sardeshmukh et al. 2000, Albers and Newman 2019, 2021)

Identifying 'forecasts of opportunity' with LIM expected skill

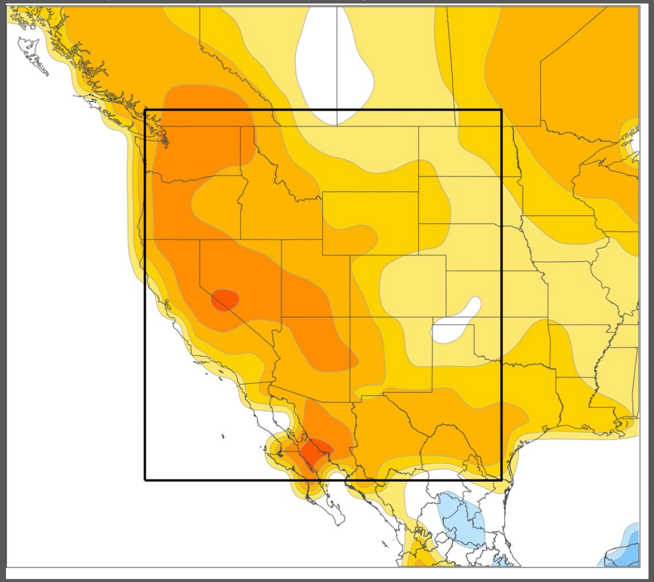
Skill of <u>all</u> Week 3-4 hindcasts April-June 1959-2021



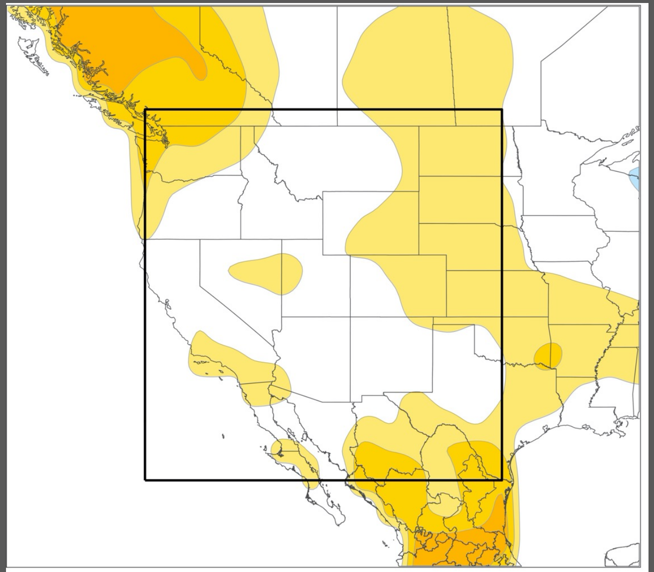
Separate hindcasts by LIM expected skill

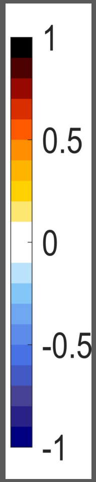
⟹

Top 10% of expected skill



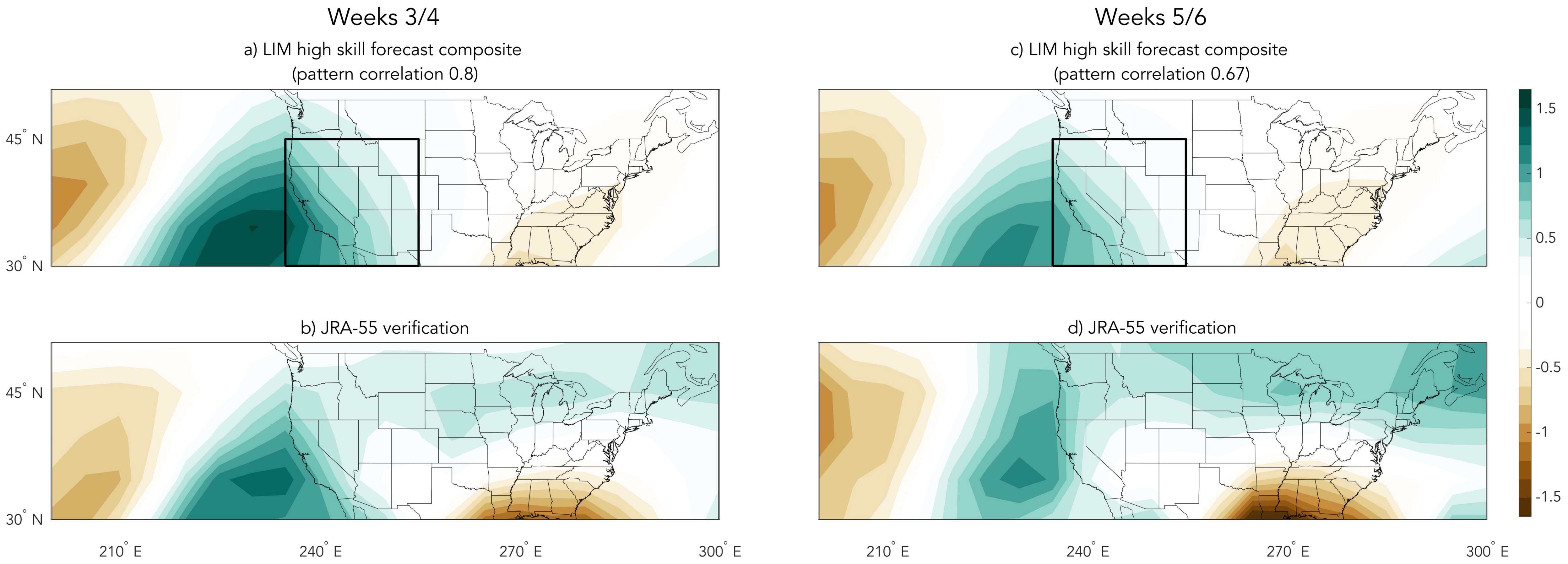
Bottom 10% of expected skill



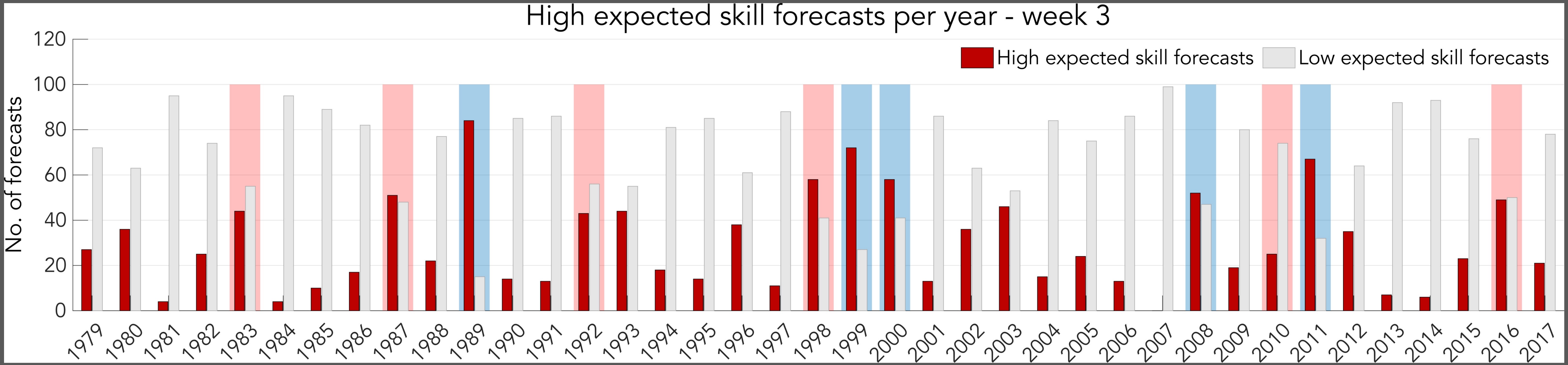


⟹ LIM successfully identifies skillful forecasts *<u>at time of forecast</u>*

Does LIM 'expected skill' help predict variables directly related to precipitation? Preliminary work suggests yes.



How does LIM expected skill compare to just using ENSO?



Conclusions:

- LIM North American 2m temperature skill is, on average, competitive with the IFS
  - IFS skill may be slightly better during spring
  - LIM skill may be slightly better during summer-fall
- Soil moisture improves forecasts over western US during early summer and Alaska and the southeastern US during late summer/early fall
- Predictable 2m temperature anomalies appear to associated with the co-evolution of soil moisture and tropical SSTs anomalies, though this linkage needs significant additional analysis
- The LIM “expected skill” can predict (at time of forecast) when North American 2m temperature forecasts will be skillful

Future directions:

- Spring transition season temperature forecasts to aid with snowpack melting predictions
- Late summer to early fall fire weather outlooks (e.g., predicting fire season ending precipitation events)
- Winter integrated precipitation/snow water equivalent predictions