Initialized seasonal to interannual forecasting without initialization

Matt Newman

University of Colorado/CIRES and NOAA/ESRL/PSD

Ding, Hui, Matthew Newman, Michael A. Alexander, and Andrew T. Wittenberg, 2018a: Skillful climate forecasts of the tropical Indo-Pacific ocean using model-analogs. *J. Climate,* doi: 10.1175/JCLI-D-17-0661.1

Ding, Hui, Matthew Newman, Michael A. Alexander, and Andrew T. Wittenberg, 2018b: Assessing ENSO seasonal forecast skill of the CMIP5 models. *Geophys. Res. Lett., submitted.*

Some known climate model forecast issues

- Model drift: model mean climate $\neq$ observed mean climate
  - *Fixes: bias correction, where mean model error as function of season and forecast lead time is removed a posteriori.*
  - *Also: flux correction, anomaly model, nudging*
- Model states $\notin$ observed phase space
  - *Fixes: model output statistics, inflation of spread*
- **Initialization shock**: initialization $\notin$ model phase space
  - *Fixes: coupled data assimilation, anomaly model*
  - This problem may be even worse for decadal forecasts where the deeper ocean initialization matters more

For forecasts made by a model that has drifted to its own climate: initialize on model’s attractor, not on nature’s attractor

"Model-analog" technique

- **Match observations to states from a long CGCM control simulation**
- Since these states are fully in balance in the model, we already know how they will evolve
- So: *construct an analog model of the model itself to make forecasts, with no additional model integration necessary (reproduce model attractor)*

- For target state: analog ensemble is the *k* nearest states, defined by root-mean-square (RMS) distance *d*
- No weighting of members: ensemble-mean forecast is mean of evolution of analog ensemble (~15-20 members from ~300-500-yr run seems sufficient)
- Analogs defined from SST/SSH anomalies from the tropical Indo-Pacific (30E-80W, 30S-30N); SST and SSH are equally weighted in *d*

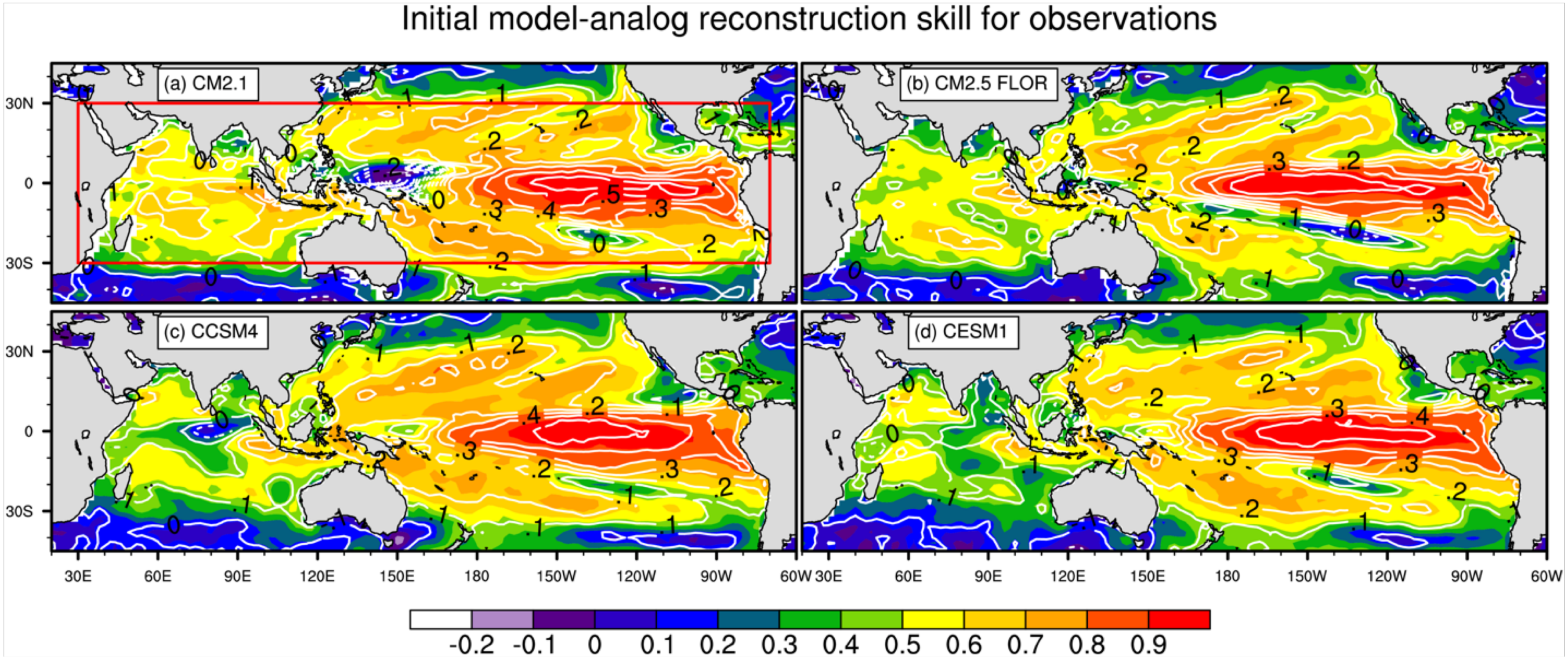
NMME models used in this analysis

Model	Year of radiative forcing	Length of run (in years)
GFDL CM2.1	1860	4000
GFDL CM2.5 FLOR	1990	700
NCAR CCSM4	1850	1100
NCAR CESM1	2000	700

Note: all models have been spun-up (no appreciable drift in tropical SST)
[This ruled out CFSv2, unfortunately, although it yielded similar results]

All models have ***fixed climate*** (i.e., fixed radiative forcing)

Initial model-analog representation of observations is only fair…



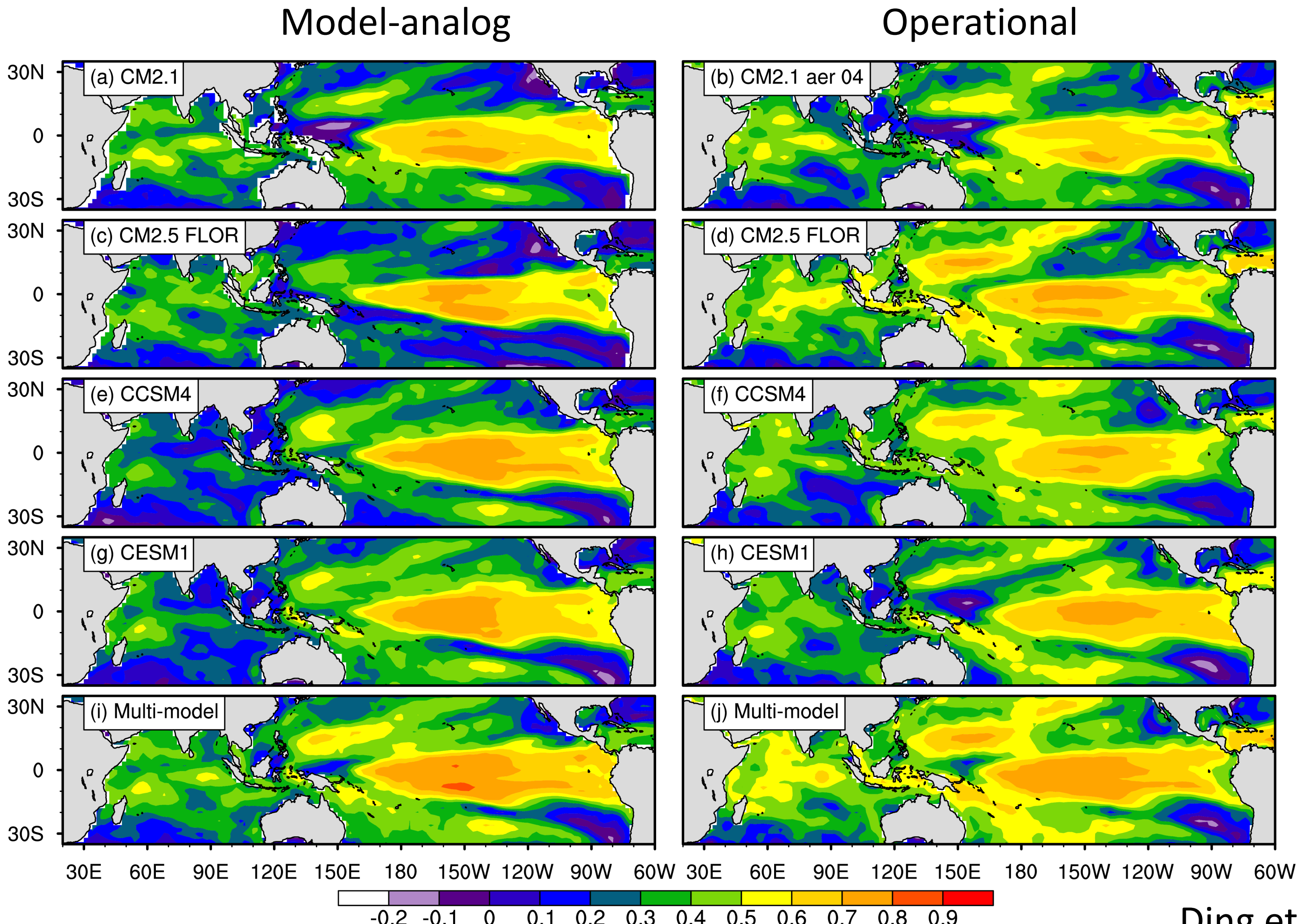
Correlation (shaded) and rms skill score (1-standardized error; contours) of ensemble mean analogs with target anomaly

Training run is entire control run for each model (varies in length)

Verification: 1982-2009 (observations)

...yet model-analog skill matches corresponding model hindcast skill (1982-2009)

Month 6 SST skill

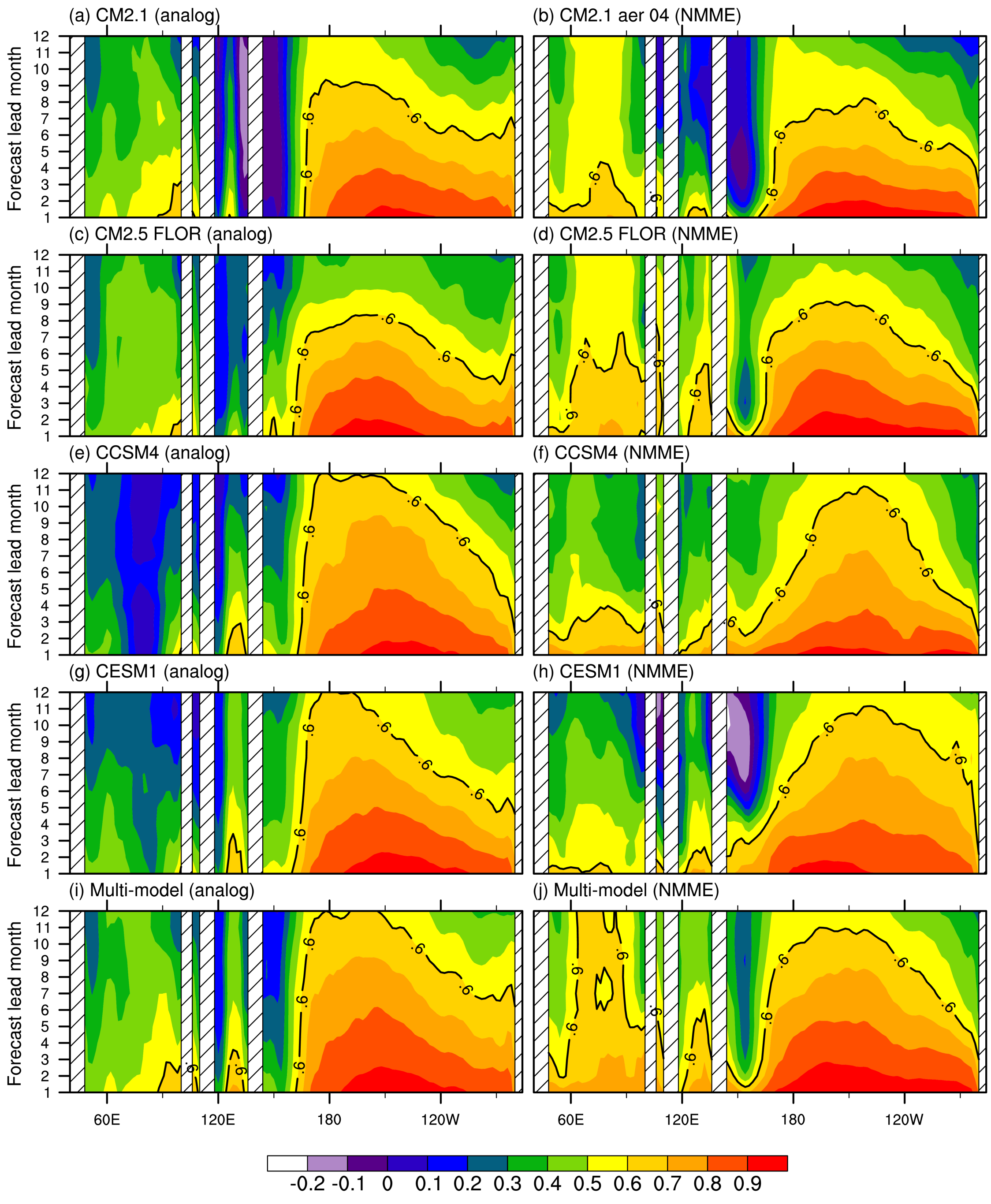


Ding et al 2018a

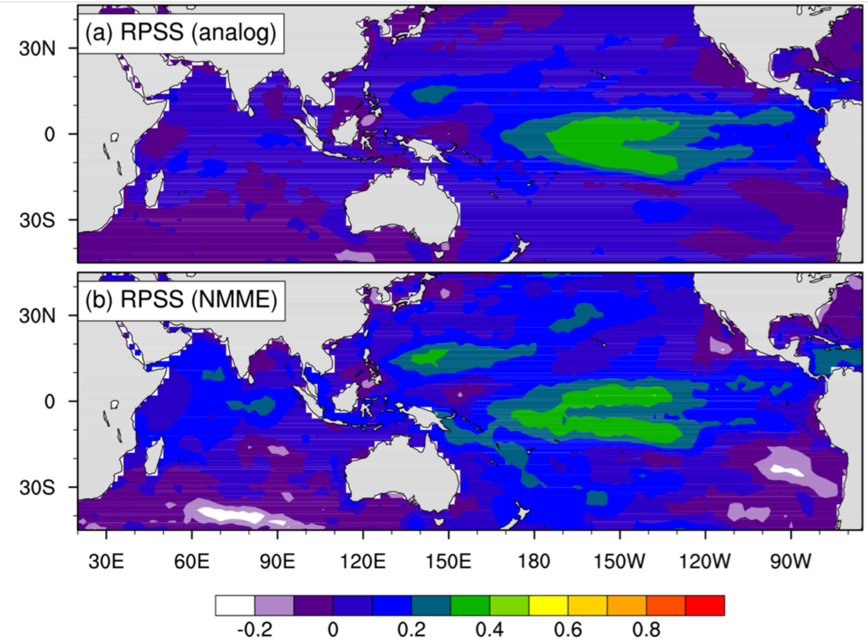
Equatorial Pacific skill: model-analog matches *or exceeds* hindcast at all forecast lead times

Equatorial rms skill score (1 – standardized error), 1982-2009

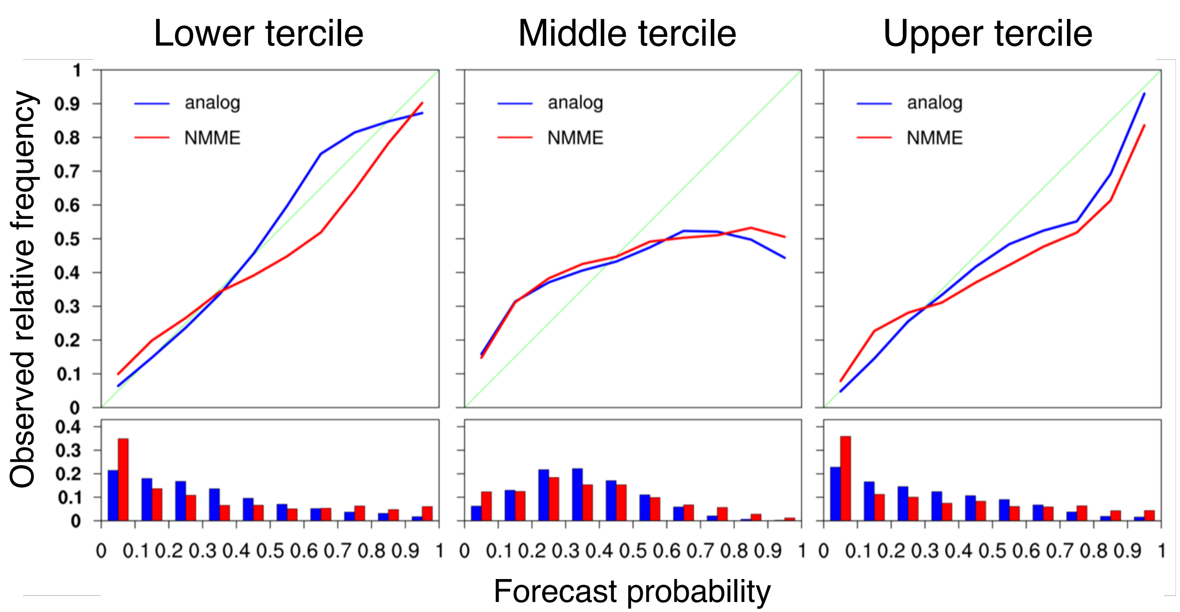
NMME hindcasts are bias corrected



Month 6 probabilistic skill: model-analog ensemble is also comparable to hindcast ensemble, *despite large initial ensemble spread*



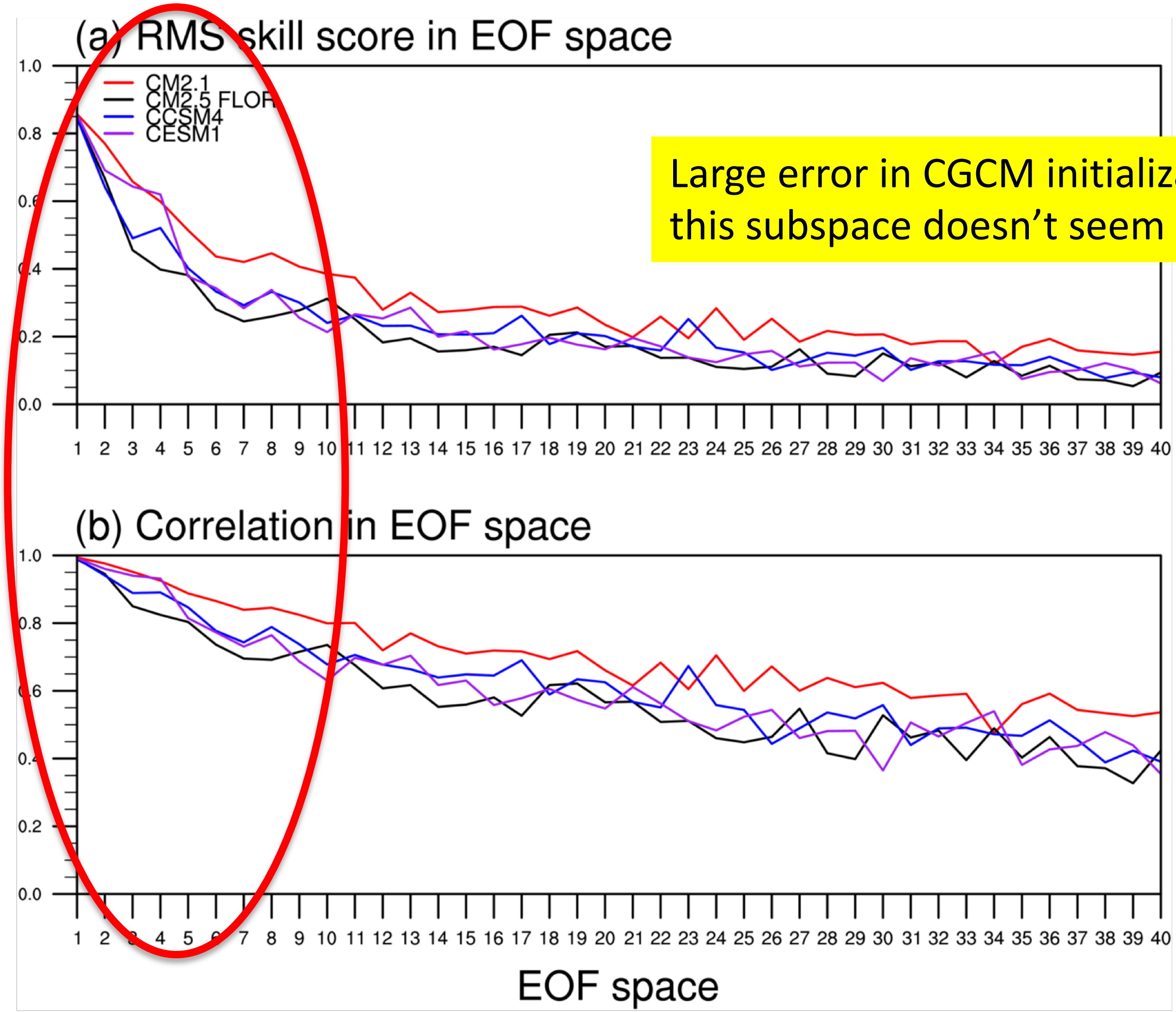
Top panels: RPSS (Rank Probability Skill Score) is higher for model-analog in tropical Pacific



Bottom panels: Reliability and frequency of occurrence (i.e., ”sharpness”): model-analogs are slightly more reliable and less sharp

Ensemble mean analog representation of target anomalies better in low order EOFs

Maybe initialization in this subspace is enough

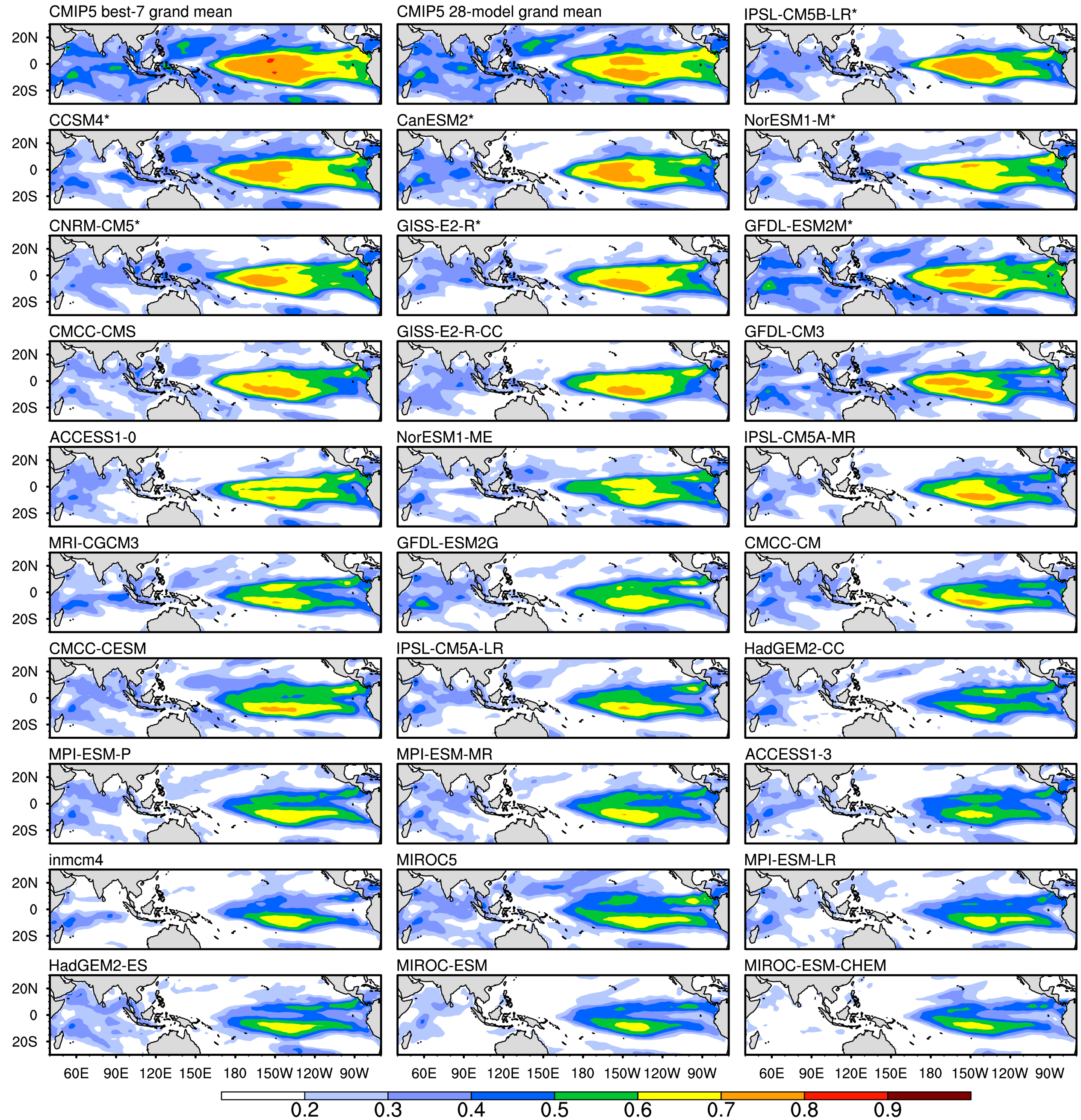


Large error in CGCM initialization in this subspace doesn't seem to matter

Turn *any* climate model into a forecast model!

Month 6 SST skill, 1982-2009

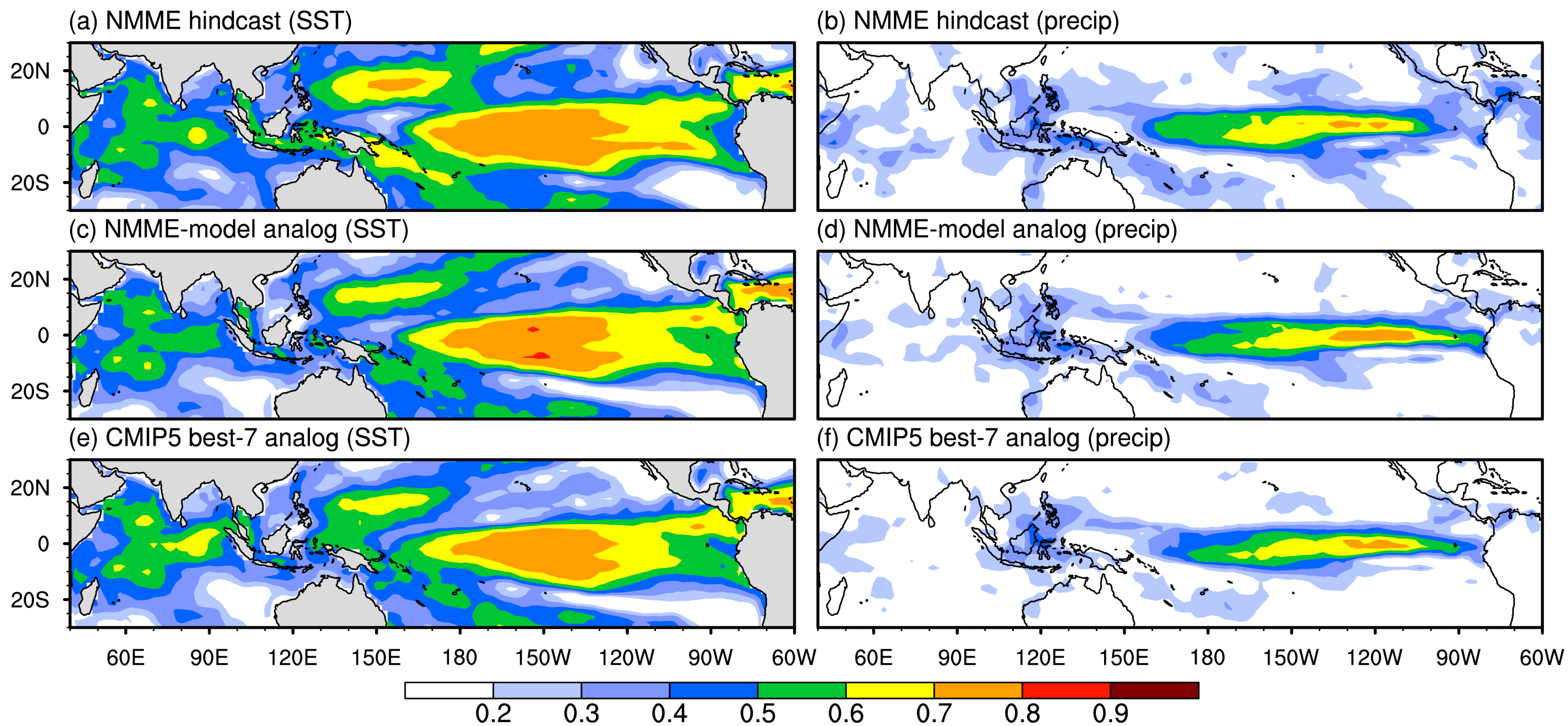
Models ordered by Month 6 Niño3.4 AC skill



Ding et al, GRL, submitted

Including CMIP5 trend forecast enhances model-analog skill in Indian Ocean

Month 6 hindcast skill, 1982-2009



Top: Operational (assimilation-initialized) NMME
Middle: NMME model-analog + trend forecast
Bottom: CMIP5 "best-7" model-analog + trend forecast
Trend forecast = CMIP5 historical run ensemble mean

Ding et al, GRL, submitted

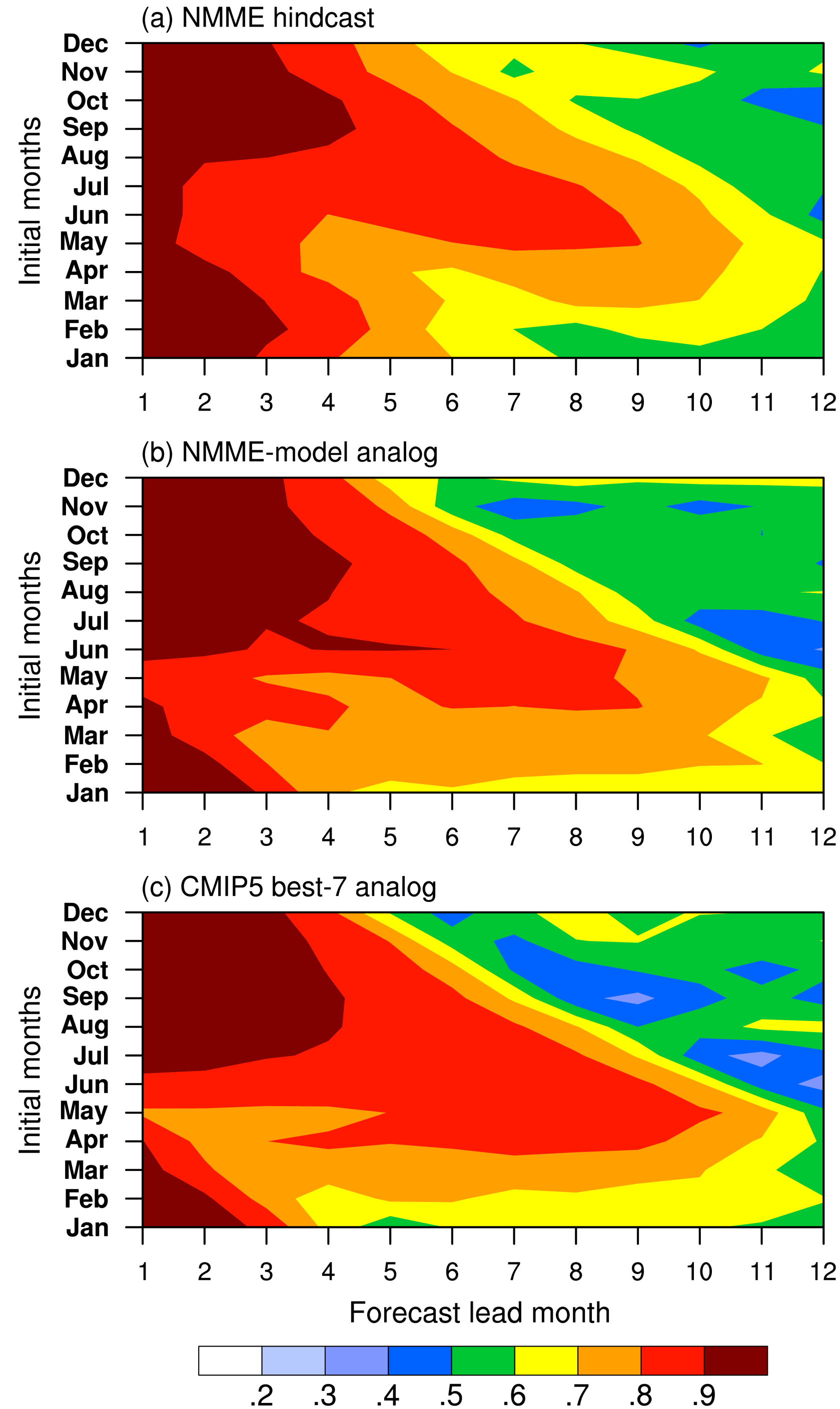
Model-analogs from both NMME and CMIP5 models capture seasonality of skill

Niño3.4 skill, 1982-2009, as a function of initialization month

Top: Operational (assimilation-initialized) NMME

Middle: NMME model-analog + trend forecast

Bottom: CMIP5 ”best-7” model-analog + trend forecast



Ding et al, GRL, submitted

Why stop with forecast leads of 12 months if it costs nothing to go longer?

Why start hindcasts in 1982 when we only need SST and SSH fields to ”initialize” the forecasts?

Niño3.4 skill

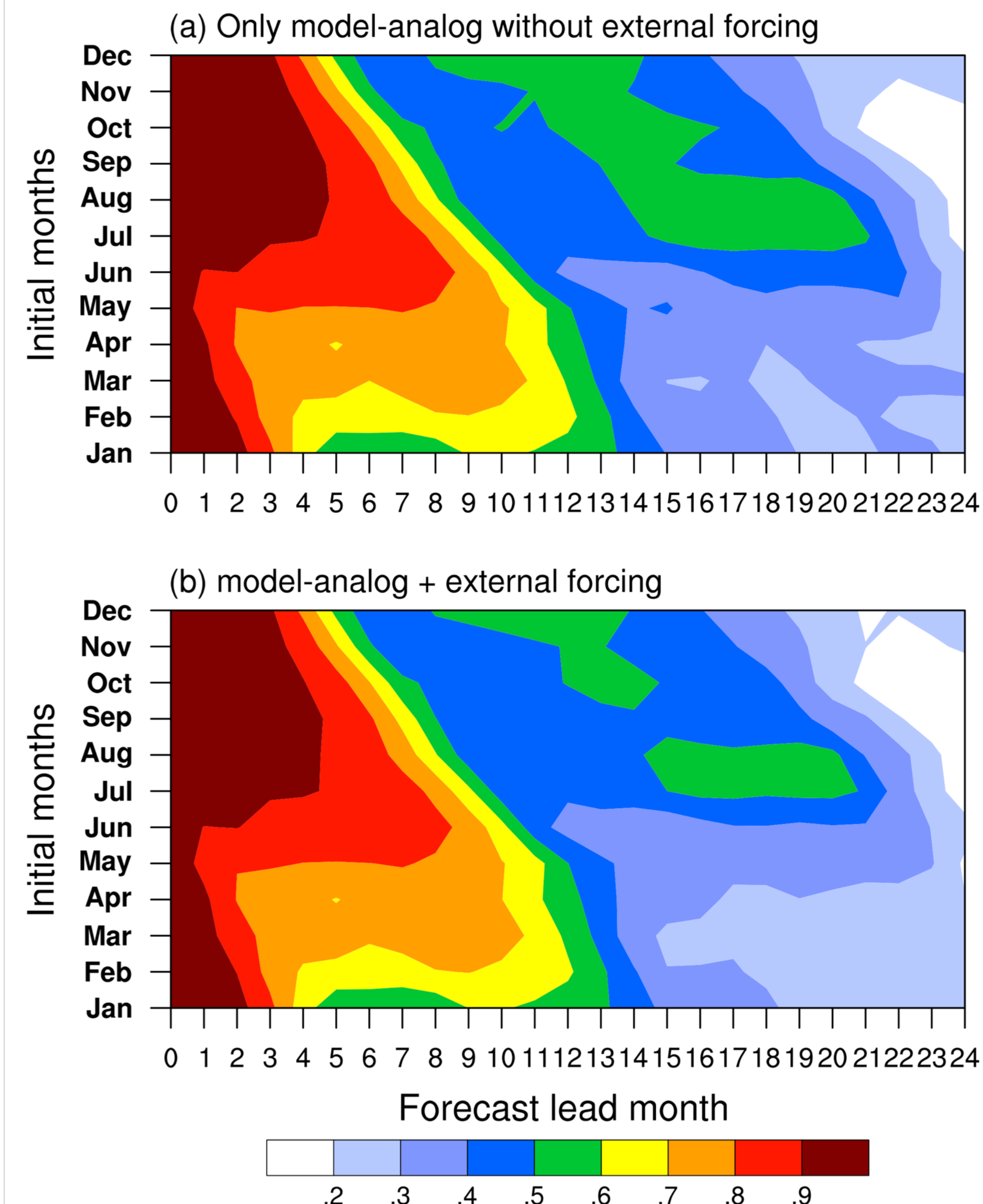
Extend hindcasts to "Year 2" (at no additional cost)

And extend hindcast database to **<u>1961-2010</u>**

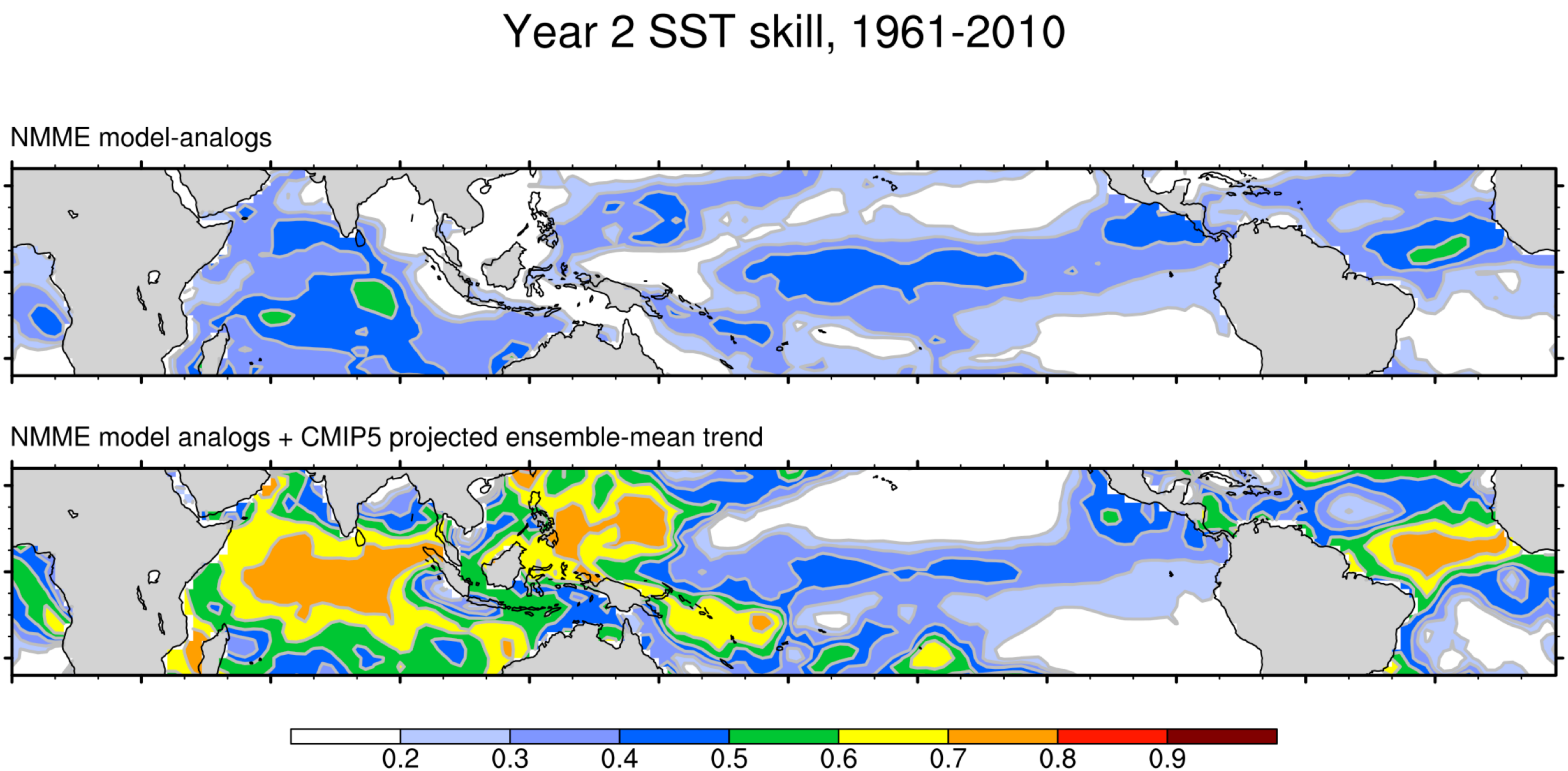
Skill shown as a function of initialization month

Top: NMME model-analog ensemble mean

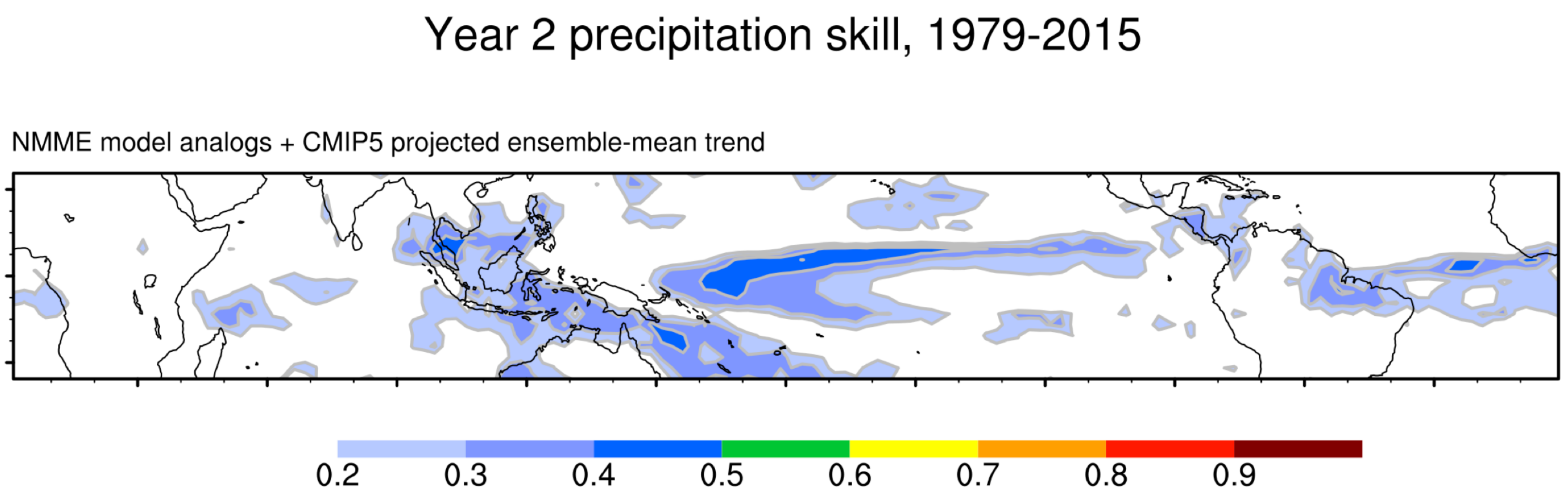
Bottom: NMME model-analog ensemble mean + CMIP5 trend forecast



Much of "Year 2" hindcast skill outside of ENSO region comes from projected trend

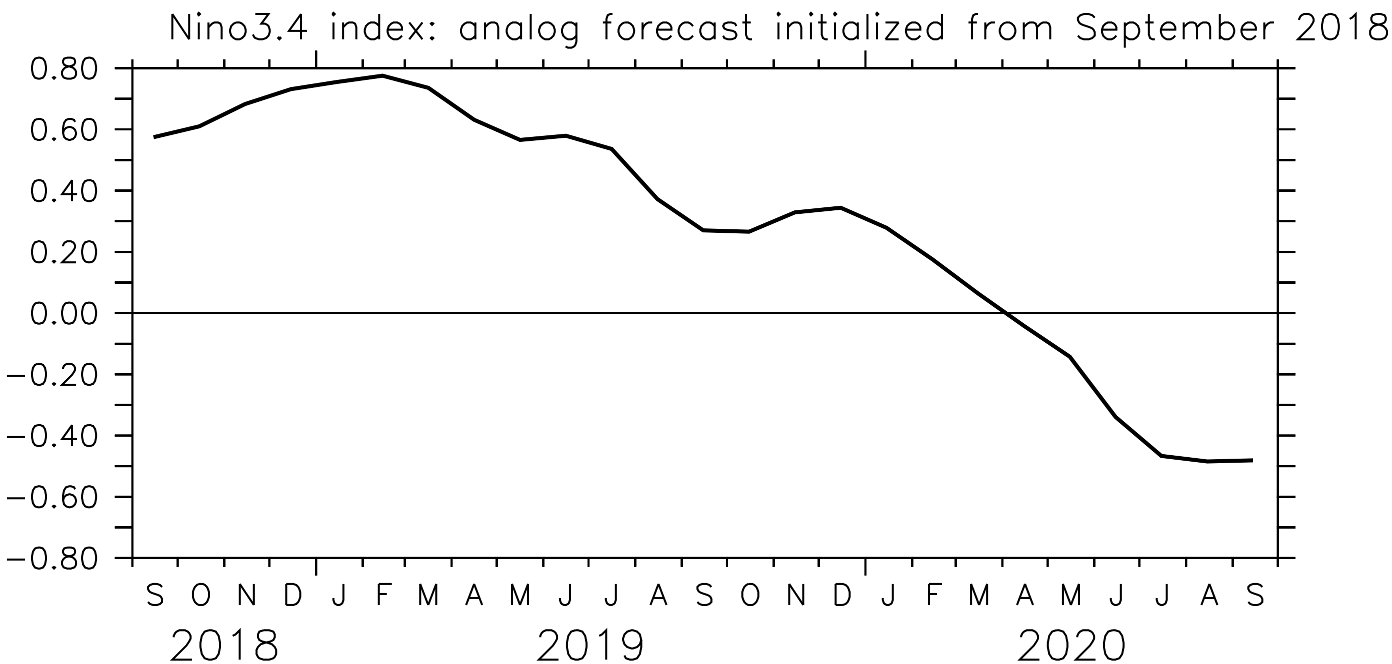


“Year 2” precipitation skill, determined from model-analog tropical Indo-Pacific SST/SSH



Current forecast: 2-yr El Niño!

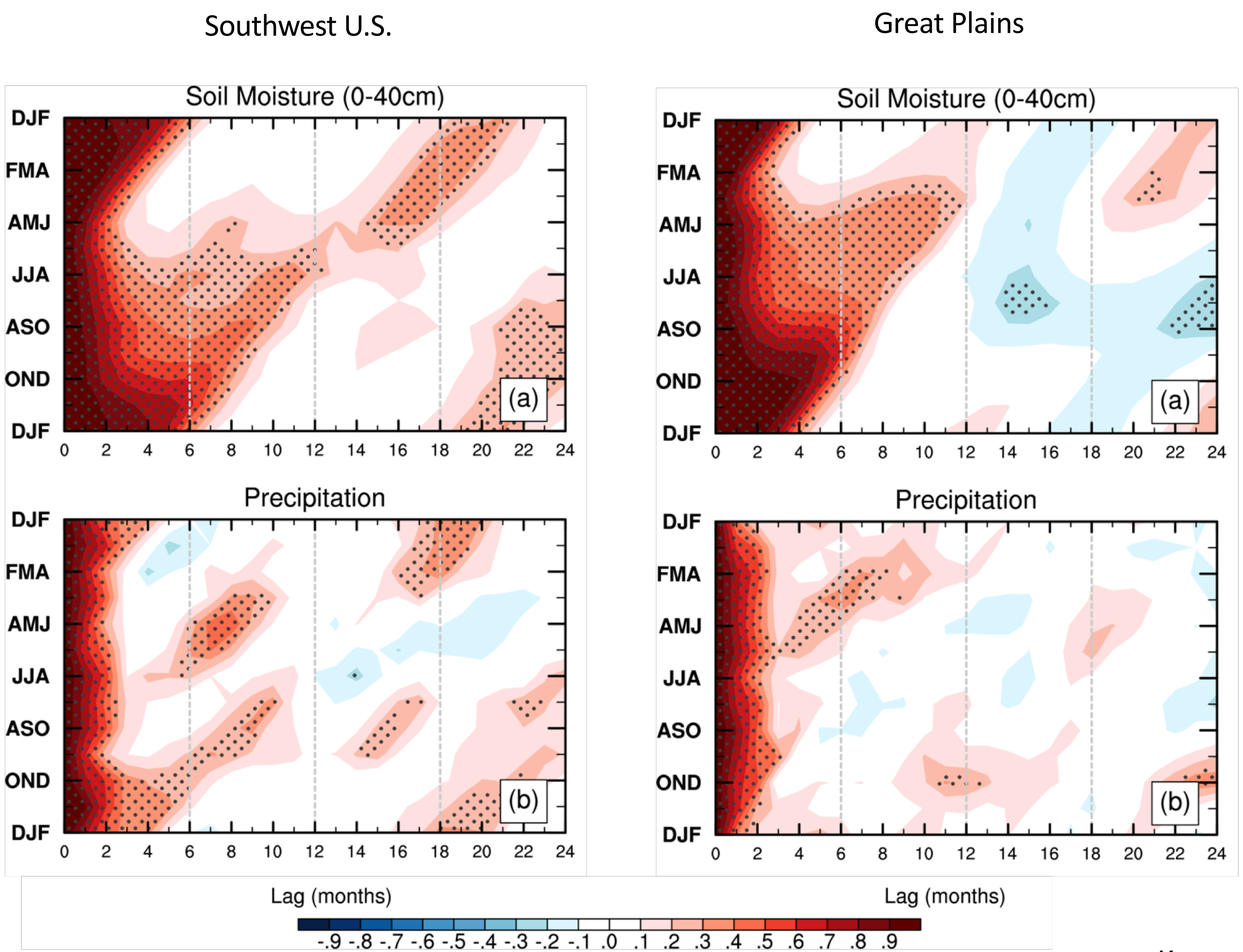
(September initialization, using GODAS SSH and ERSST.v5 SST, detrended)



Conclusions

- Model-analogs can match forecast skill of traditionally-initialized CGCM forecast models of monthly tropical Indo-Pacific SST/SSH/precipitation anomalies
- Where model-analog skill is higher, is it due to:
  - Lack of initialization shock?
  - Better bias correction?
  - Control runs that do not have erroneous regional trends?
- Initialization may only need to be accurate within low-order, large-scale subspace
- Any (good) CGCM with a (sufficiently long) control run can produce skillful monthly forecasts
  - Evaluate model-analogs as part of model development
  - Large model ensembles → more samples for weighting/calibration
- **Model-analogs allow easy extension of forecast leads and hindcast datasets → develop "Year 2" forecasts**

Root zone soil moisture autocorrelation (1950-2010) from CLM "LDAS" dataset also suggests Year 2 predictability, although only in some seasons



Kumar et al. 2018

Skill of low-order linear empirical dynamical model is comparable to NMME ensemble mean

Linear Inverse Model (LIM) and NMME mean have similar patterns of skill, *which can be explained by <u>expected LIM skill</u>*

LIM: $d\mathbf{x}/dt = \mathbf{L}\mathbf{x} + \mathbf{F}_s$

$\mathbf{x}(t)$: series of maps, $\mathbf{L}$: stable operator, $\mathbf{F}_s$: white noise (also maps)

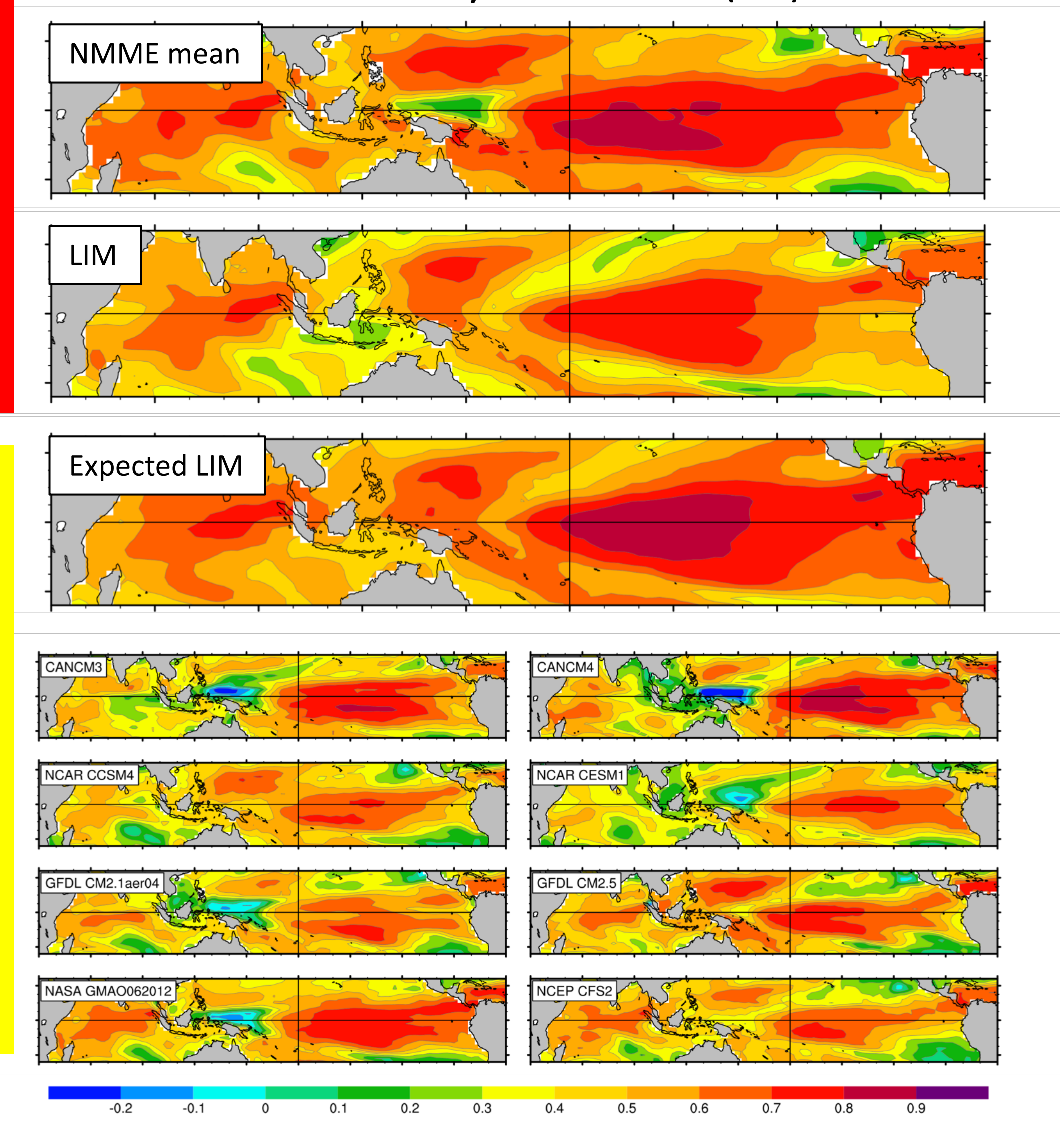
* Determine **L** from data: multivariate lag-1 autocorrelation of monthly SST/SSH/wind anomalies (reduced EOF space)

*. (Ensemble mean) forecasts for lead **τ** :

$$\mathbf{x}(t+\tau) = exp(\mathbf{L}\tau)\mathbf{x}(t)$$

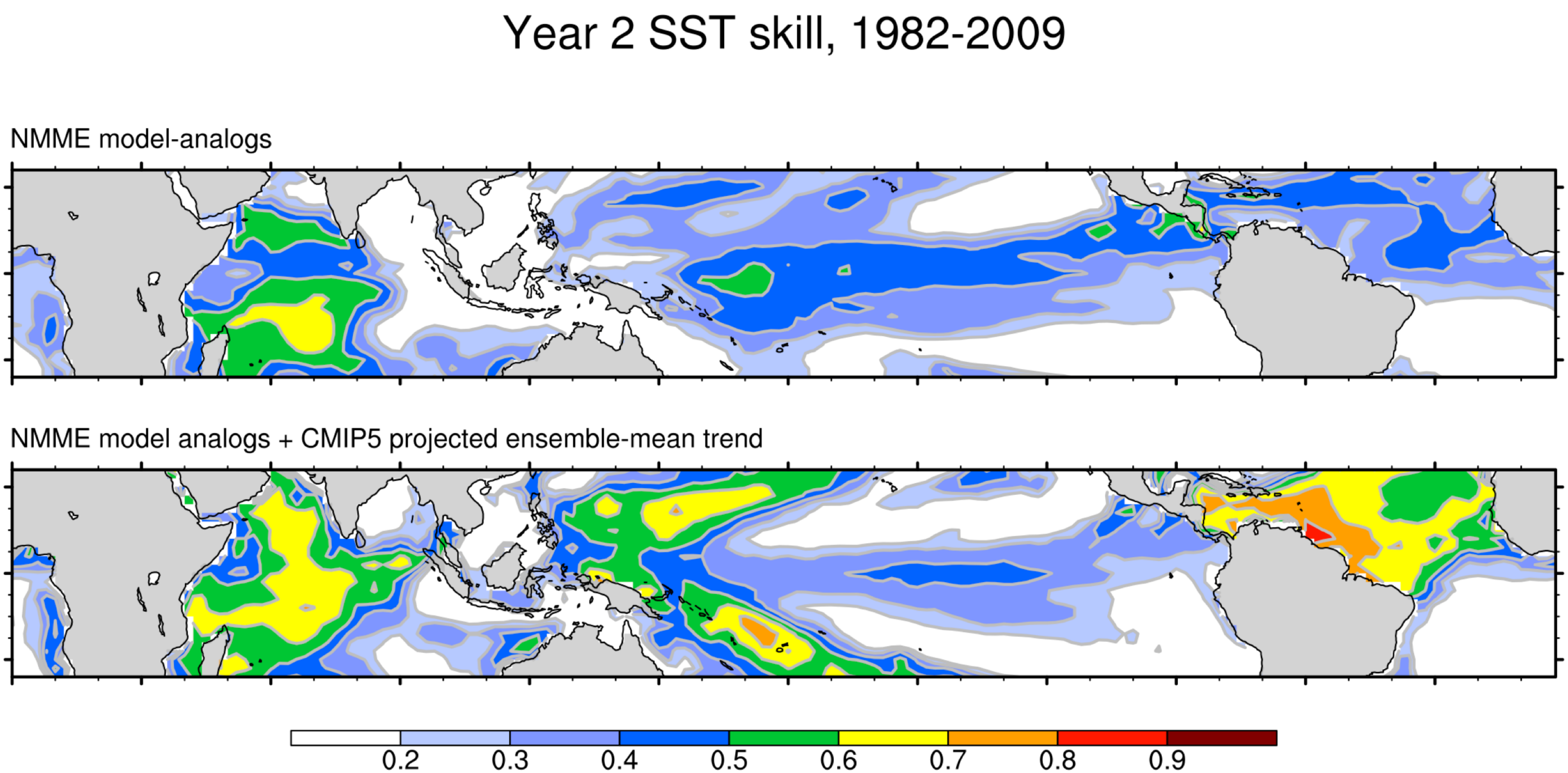
* Expected skill: *f(forecast signal-to-noise)*

Month 6 anomaly correlation (AC) skill



Newman and Sardeshmukh 2017, GRL

Recent period has somewhat higher skill

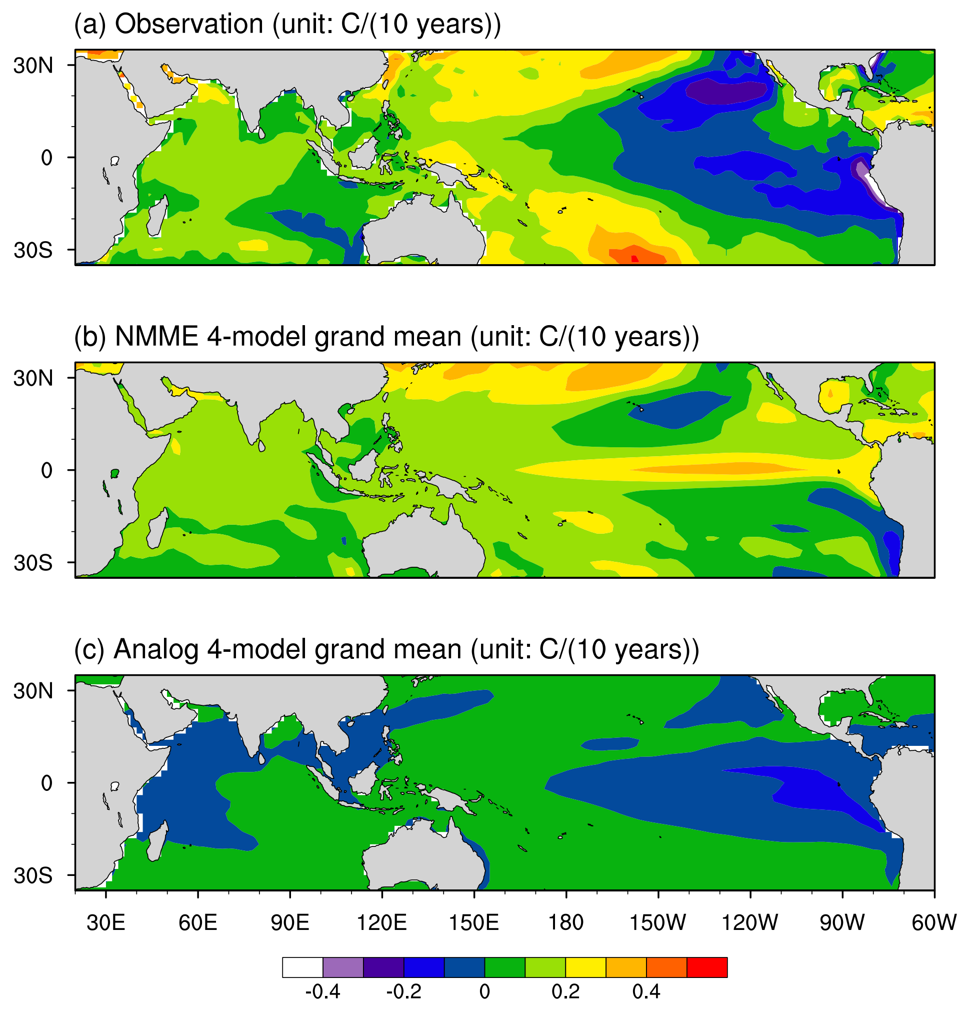


How much of the model-analog skill is due to getting the trend right or wrong?

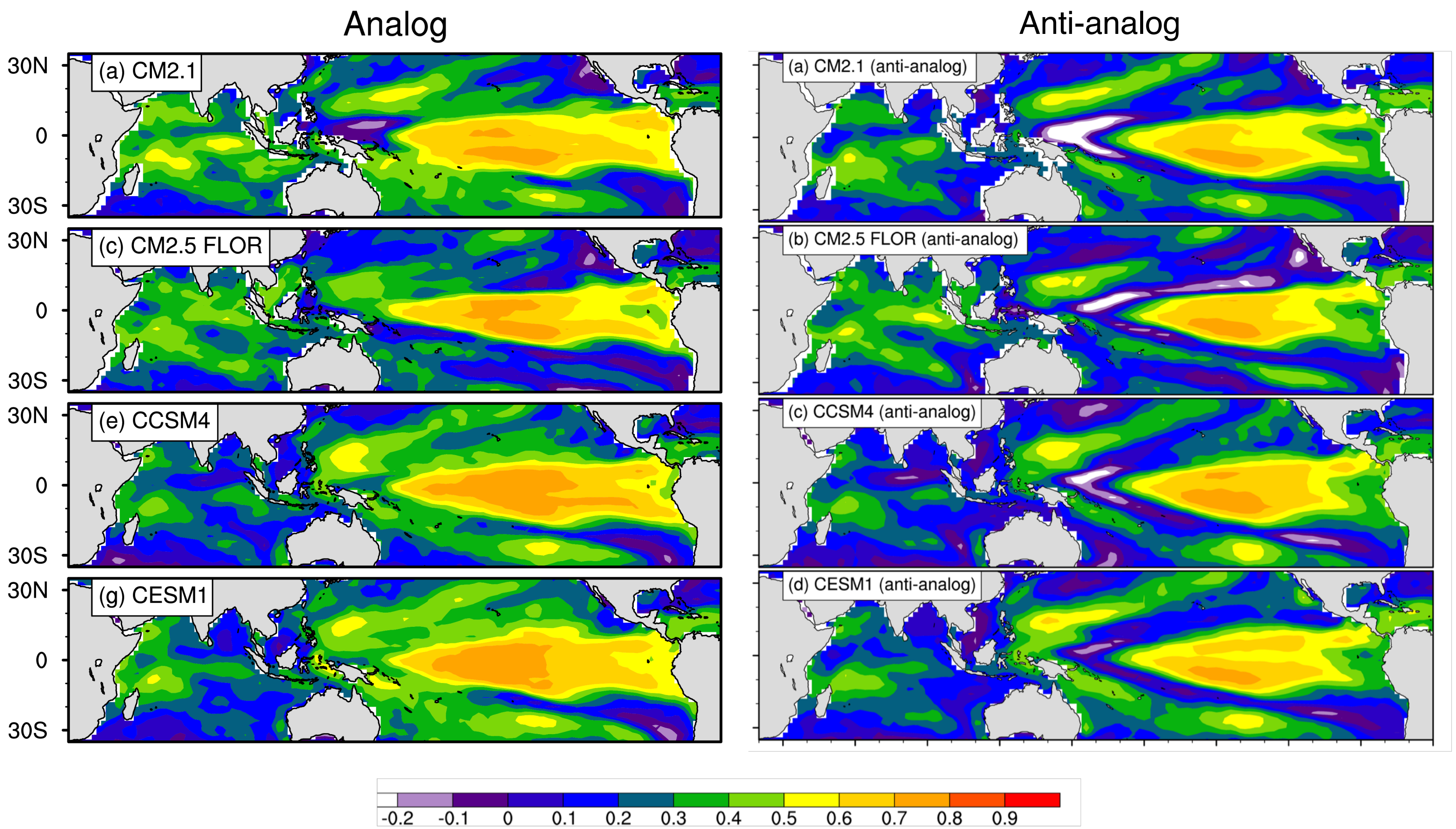
Model-analog hindcasts do not capture

- Observed Indian Ocean /warm pool warming trend, which may degrade model-analog skill
- Erroneous equatorial eastern Pacific warming trend, which may degrade NMME hindcast skill

Observed and Month 6 1982-2009 SST trends



Much of the model-analog skill is linear



Anti-analog: same as model-analog but change sign of target first

Where skill is similar, initial sign didn't matter → linear skill

Testing sensitivity of ensemble size and data library length, 1982-2009