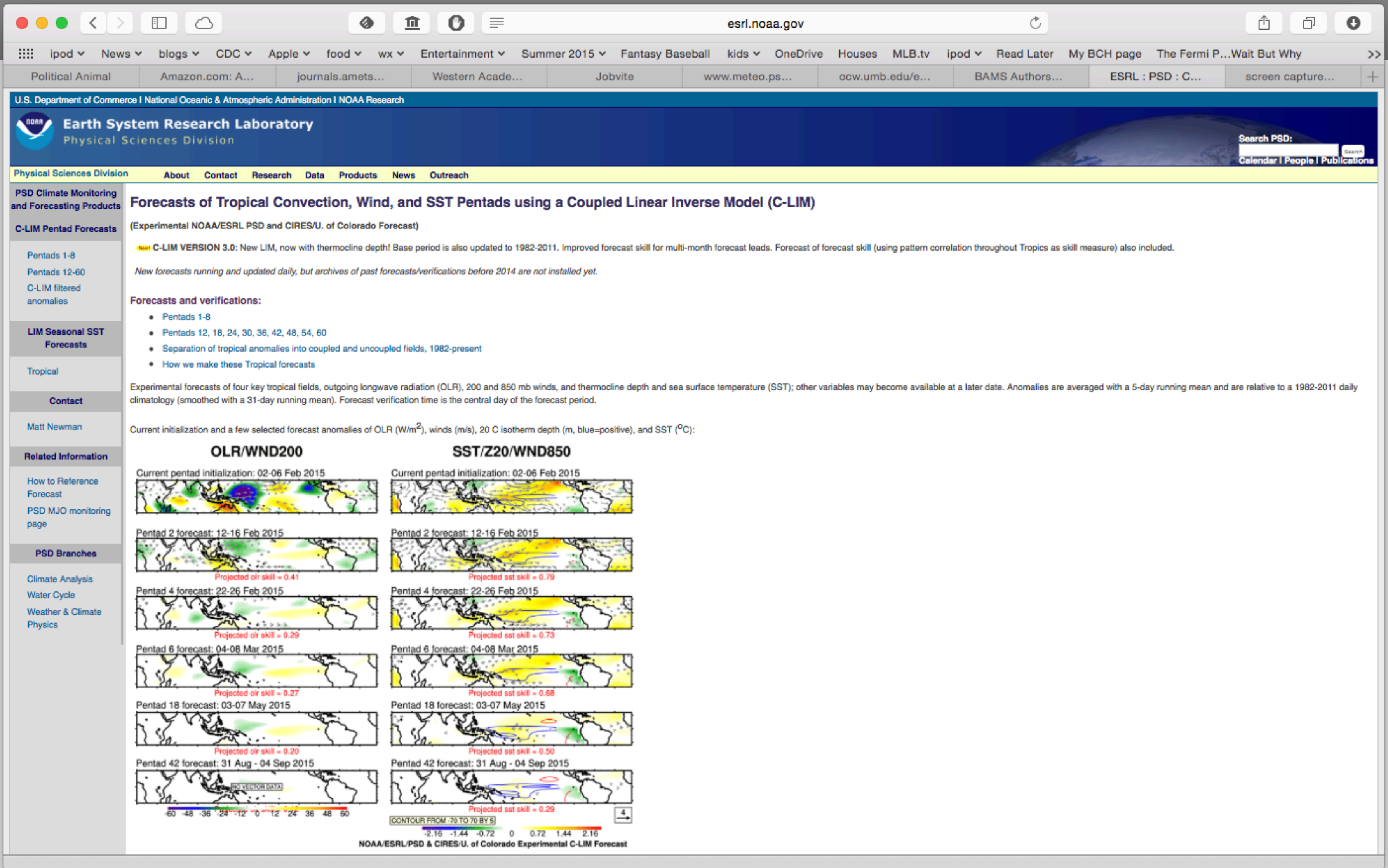
Diagnosing ENSO predictability from observations and models

Matt Newman
CIRES, University of Colorado and NOAA/ESRL/PSD



“Multivariate Red Noise*” null hypothesis

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

$\mathbf{x}(t)$ is a series of maps, **L** is stable, and $\mathbf{F}_s$ is white noise (maps)

- Determine **L** and $\mathbf{F}_s$ using “Linear Inverse Model” (LIM)
  - **x** is ocean (***SST/Z20***) and atmosphere (**OLR/*200&850 mb wind***) 5-day running mean anomalies in Tropics, 1982-2011 (similar to Newman, Sardeshmukh, Penland 2009 and Newman, Alexander, Scott 2011)
  - prefiltered in reduced EOF space
  - LIM determined from specified lag τ_o=5 days (e.g., the data averaging interval) as in AR1 model, using τ_o- and zero-lag covariance of **x**
  - Test the LIM over much longer time intervals: observed spatio-temporal lag-covariance statistics very well reproduced
  - Hindcasts determined from cross-validation (10% data withheld to recompute **L**)

*Multivariate Ornstein-Uhlenbeck process

Using LIM to estimate predictability

$$\mathbf{d\,x/dt} = \mathbf{L\,x} + \mathbf{F_s}$$

$\mathbf{L}$ = constant, $\mathbf{F_s}$ = additive (state-independent) noise.

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

"signal"

"noise"

Expected forecast error covariance

(assuming no initial error) :

$$\mathbf{E}(\tau) = \langle \varepsilon\varepsilon^{\mathrm{T}} \rangle = \mathbf{C}(0) - \mathbf{G\,C}(0)\,\mathbf{G}^{\mathrm{T}}$$

Expected forecast anomaly correlation

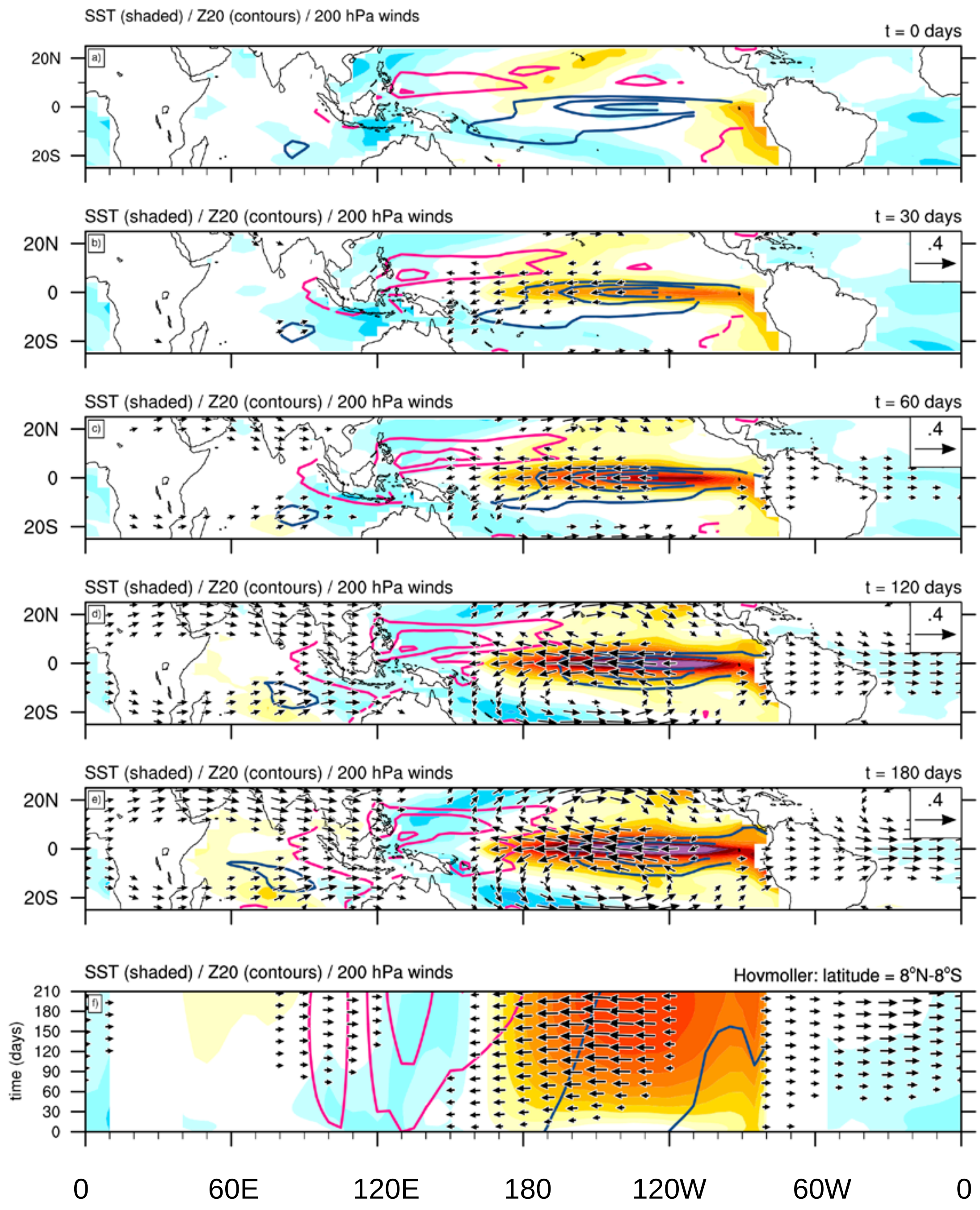
$$\rho_\infty = \frac{s}{\sqrt{1+s^2}}\text{, where } s^2 = \frac{[\mathbf{G\,C}(0)\,\mathbf{G}^{\mathrm{T}}]_{ii}}{[\mathbf{E}(\tau)]_{ii}}$$

Larger signal related to leading singular vector of G(τ)

Assumption (overly restrictive?): E(τ) depends *only* on τ

“Optimal” structure leading to greatest tropical SST anomaly growth over 180 days

Shading: SST
Contours: Z_{20}
Vectors: 200 mb winds



Compare LIM to (some of the) National Multi-model Ensemble (NMME) Hindcasts, 1982-2010

LIM hindcasts averaged into 35-day "months"

so "Month 6" skill is days 146-180

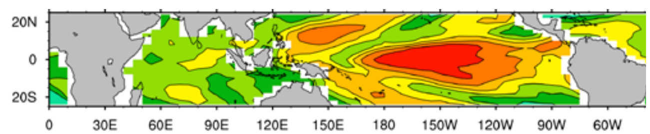
NMME hindcast ensemble monthly means from

- NCEP CFSv2
- NASA GEOS5
- NCAR CCSM3
- GFDL CM2.1/2.5
- CMC1-CanCM3/4
- ECHAM 4.5

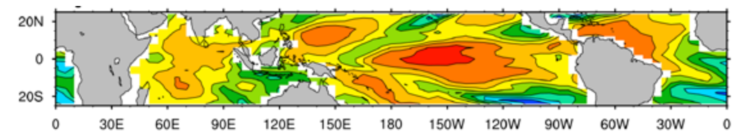
"Month 0.5" skill is first month, initialized in mid-month (see Kirtman et al. 2014)

Comparing predicted and actual forecast skill, or: some places are more predictable than others

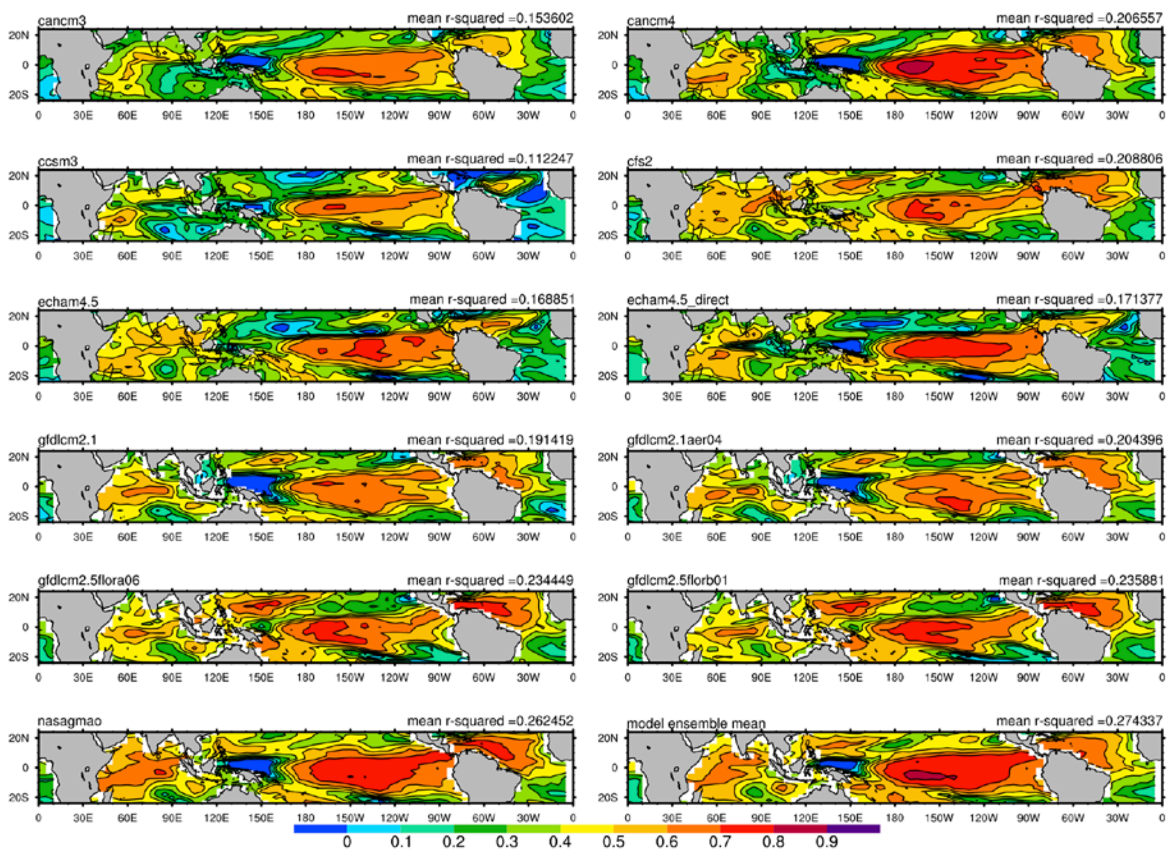
Predicted Month 6 LIM skill



Actual Month 6 LIM skill



NMME



NMME ensemble mean

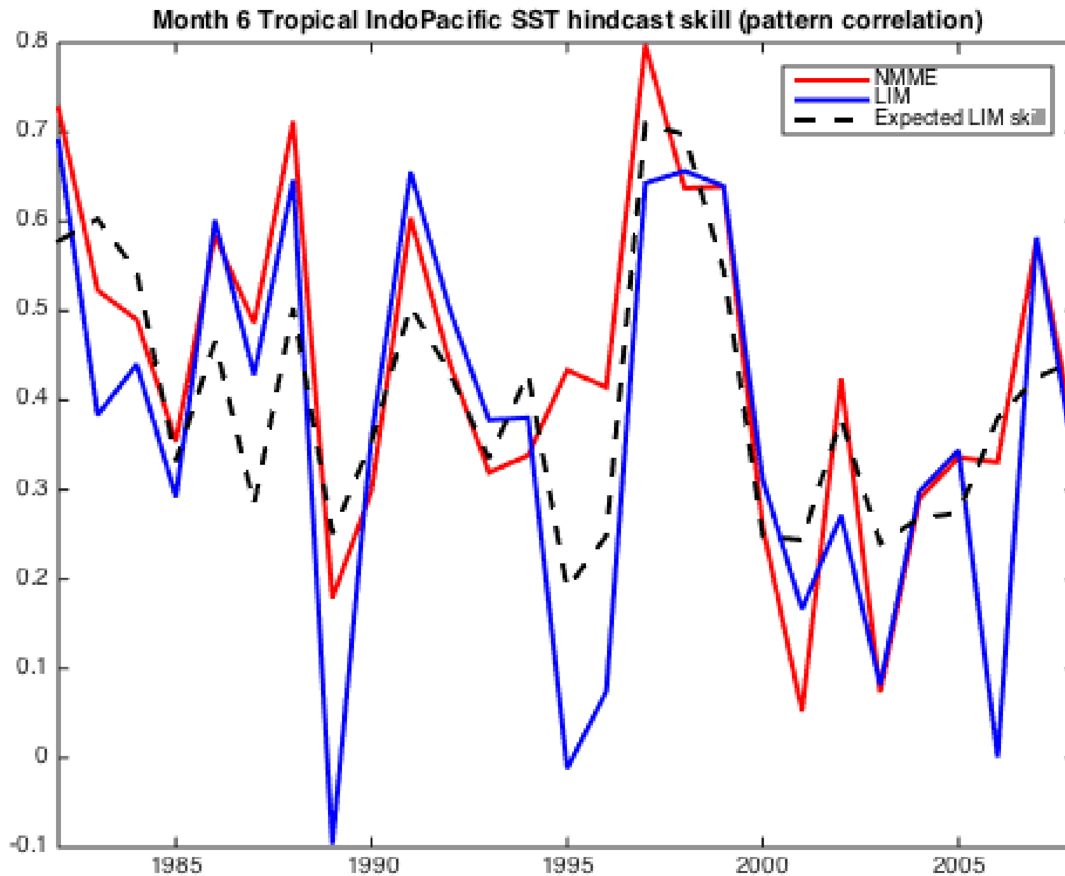
Comparing predicted and actual forecast skill, or: some years are more predictable than others

Skill averaged by year (based on initialization date)

Actual skill: Pattern correlation of tropical SST anomaly with SST forecast anomaly

Blue: LIM (0.40)
Red: NMME (0.45)

Predicted: ρ_∞ (0.43)



NMME and LIM skill correlated with ρ_∞ at 0.8

A few other points

- LIM useful for diagnosis of predictability, ***because*** its forecast skill is comparable with coupled GCMs and it reproduces observed spatio-temporal statistics
- Subseasonal-interannual forecast skill may itself be *predicted* based on LIM signal-to-noise
  - In LIM, there is no “spread/skill” relationship, but note this need not be a constraint for all linearly predictable systems
- Year-to-year variations in forecast skill in the last few decades may be due to random variations in initial conditions and not necessarily to long-term “base state” changes

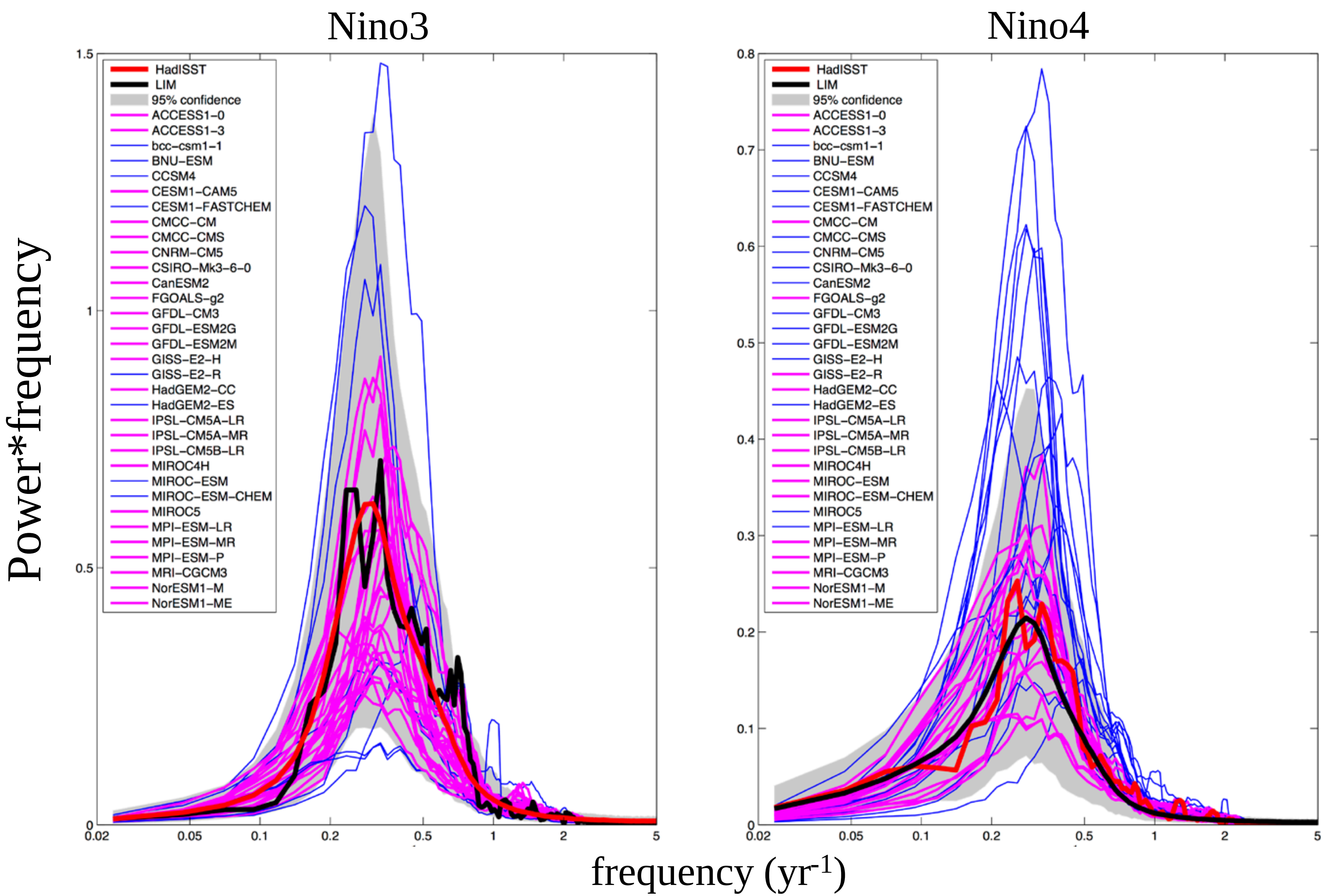
But…

what if there is a base state change?

To answer,

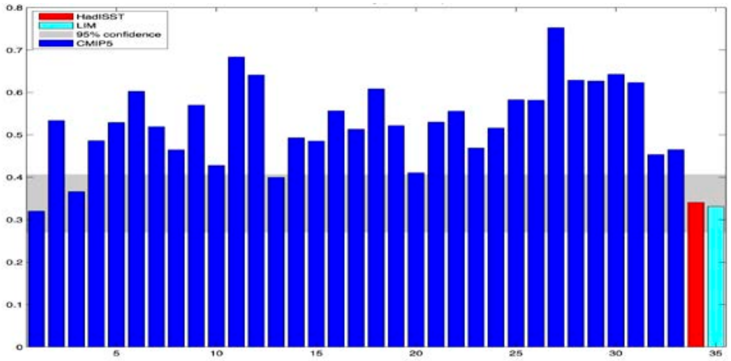
travel to other M-class planets

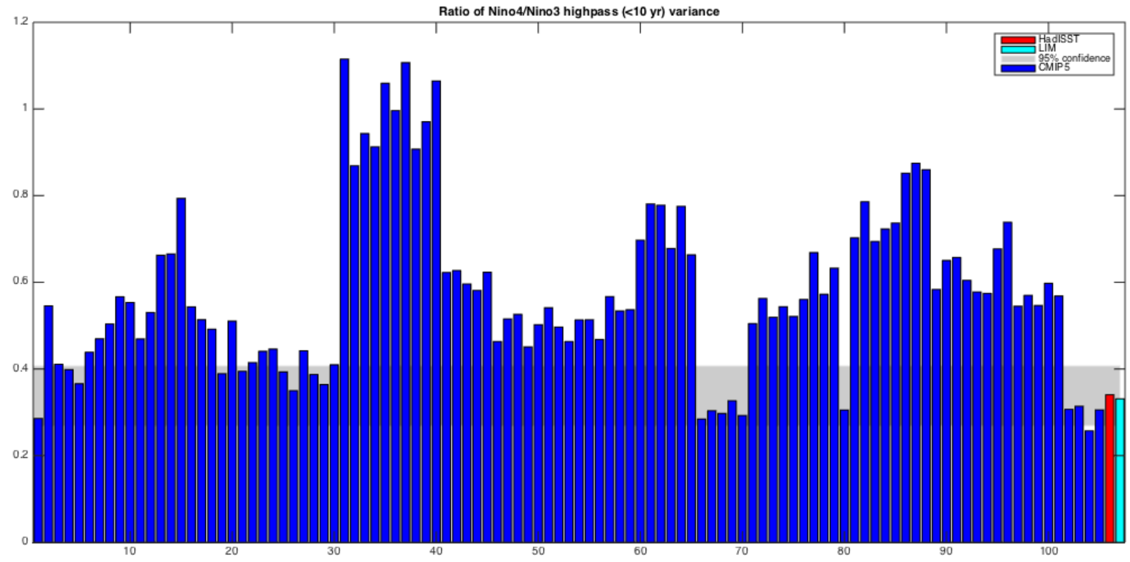
LIM based on one-season lag (now using **SST, Z20, zonal wind stress**) reproduces observed spectra; use for significance testing (blue: significantly different)



LIM based on one-season lag reproduces observed spectra and can be used for significance testing
CMIP5: Nino4 significantly too strong compared to Nino3

Ratio of Nino4/Nino3 highpass (<10yr)



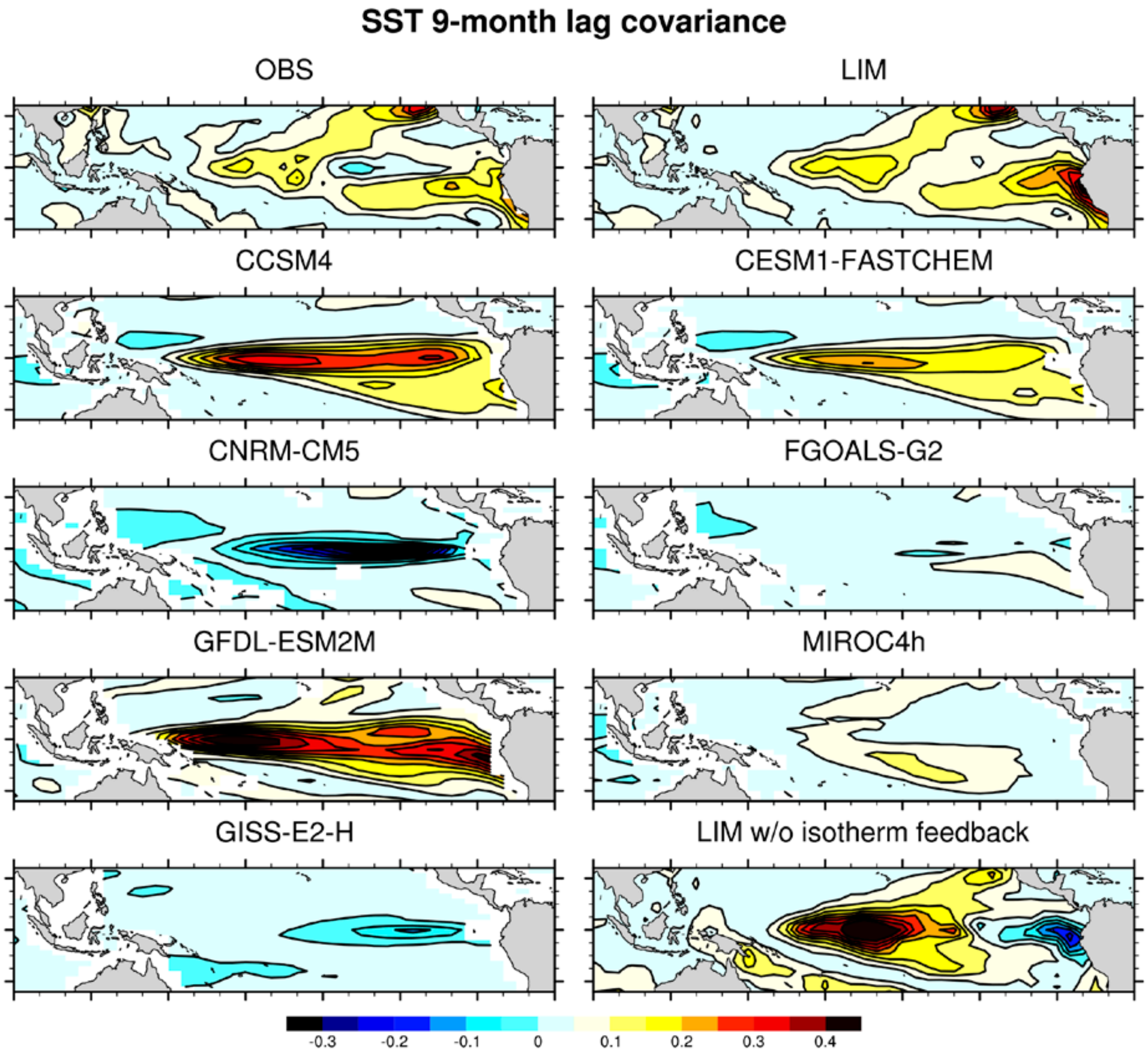


Control

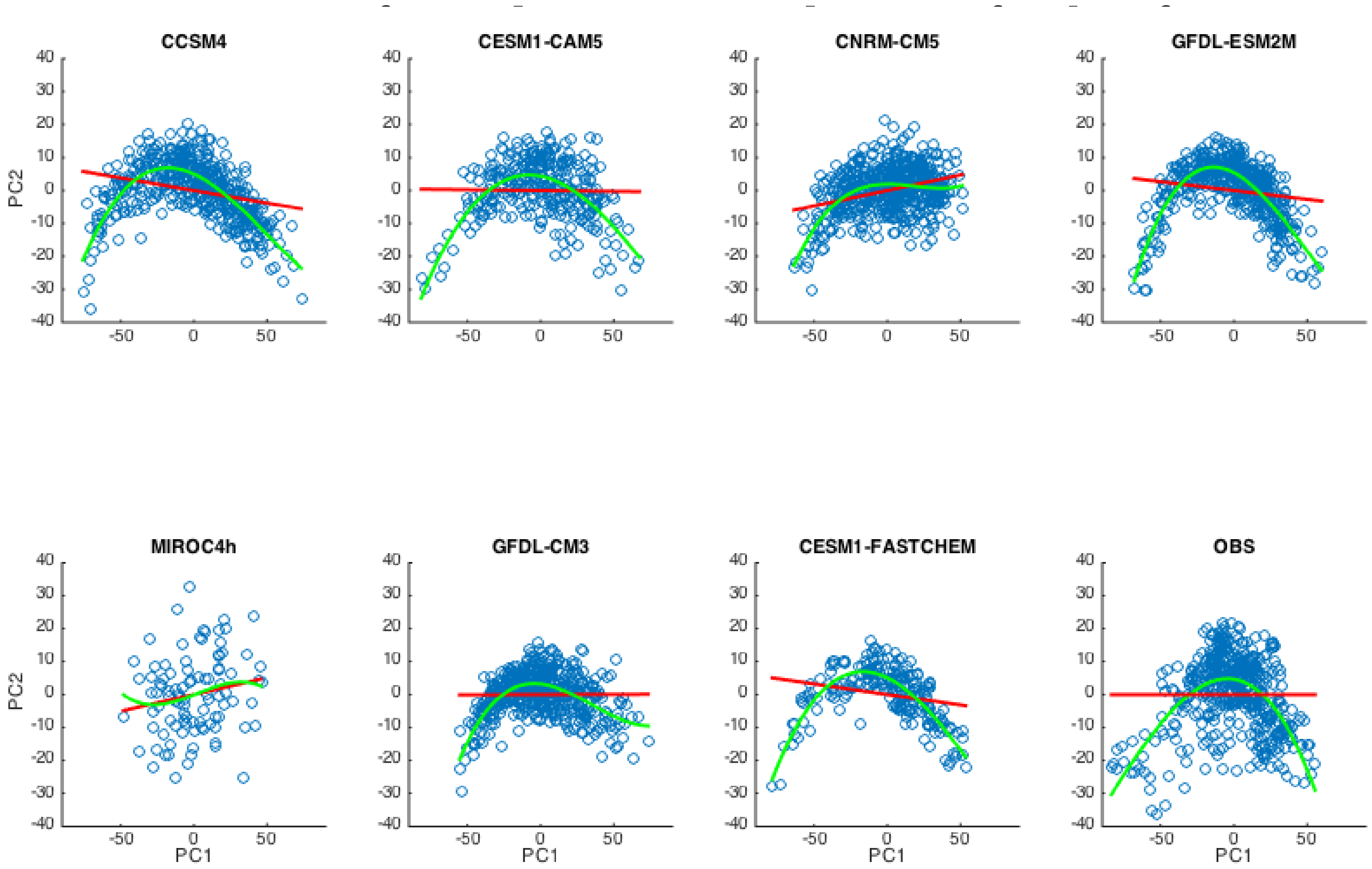
Historical

CMIP5 model index

LIM based on 1-season lag reproduces evolution statistics at much longer leads.



How nonlinear is PC1/2 in the models compared to observations?



How nonlinear is ENSO in models compared to obs? ***CMIP5: nonlinearity in PC1/PC2 plane depends on ENSO strength***

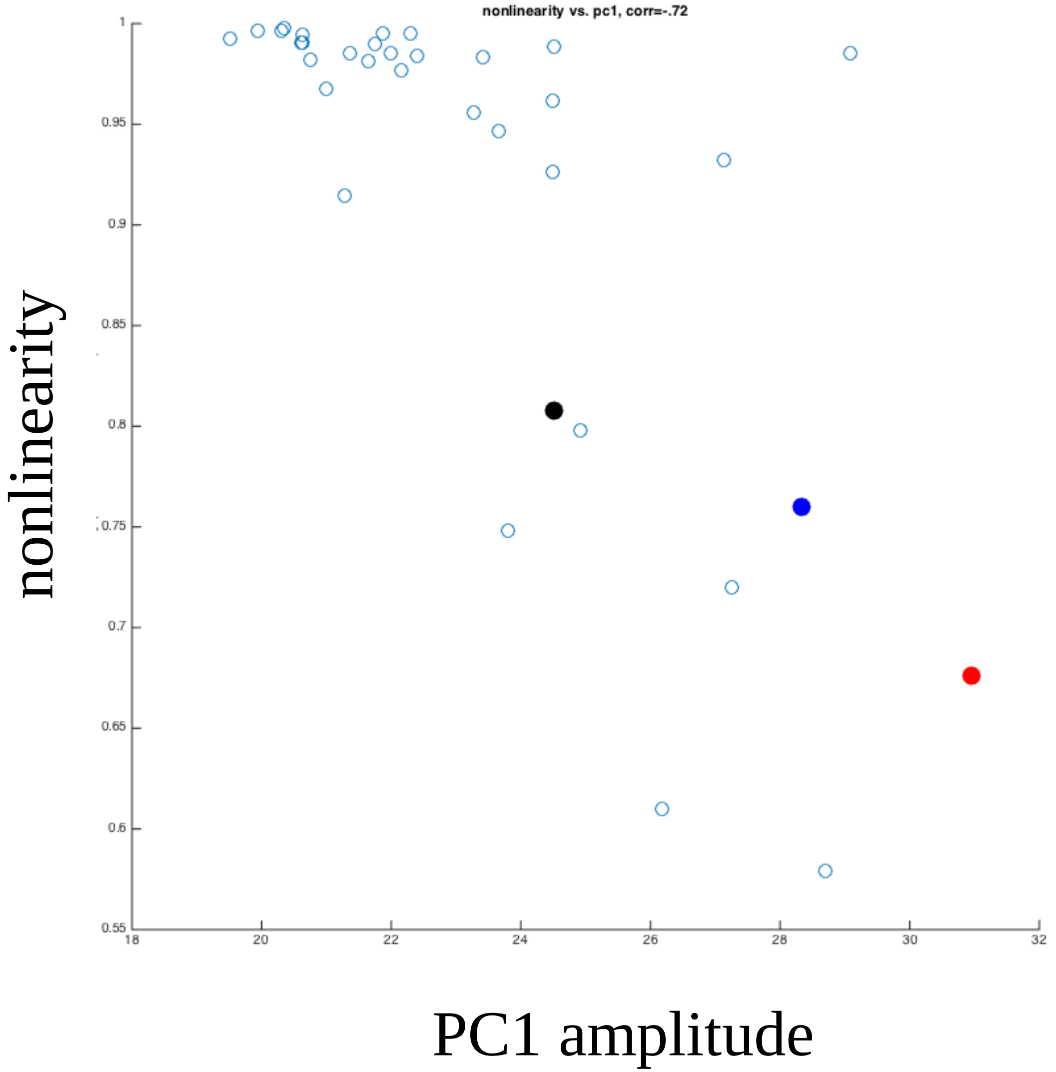
Measure of "nonlinearity":

$$\frac{\text{std error of cubic fit}}{\text{std error of linear fit}}$$

OBS

CCSM4

CESM1-CAM5



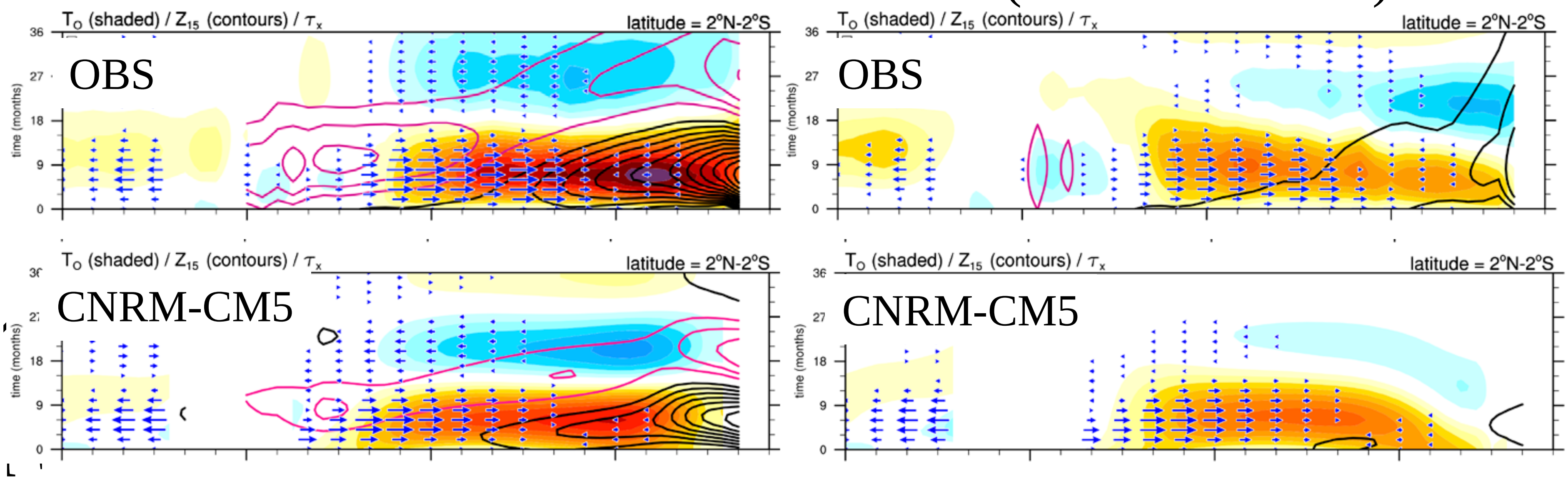
Evolution of 6-month optimal structure

SST: shading
Thermocline depth: contours
Zonal wind stress: vectors

Equatorial Hovmuller, time running from 0 to 36 months

FULL

No thermocline term
("Slab-like" LIM)



Thermocline term *strengthens* ENSO in east but slightly damps in central Pacific

Evolution of 6-month optimal structure

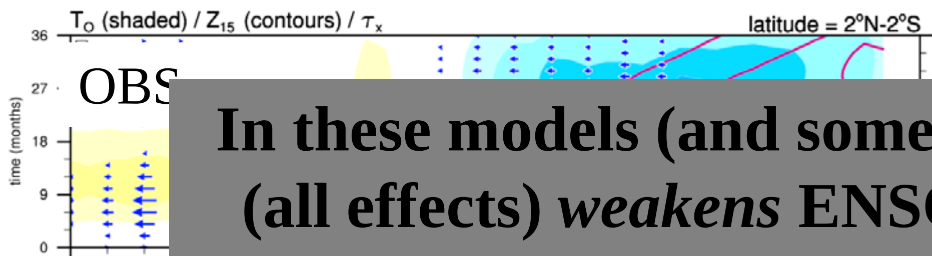
SST: shading
Thermocline depth: contours
Zonal wind stress: vectors

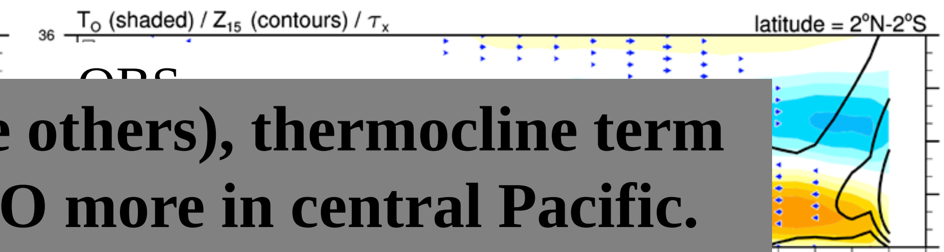
Equatorial Hovmuller, time running from 0 to 36 months

FULL

No thermocline term
("Slab-like" LIM)

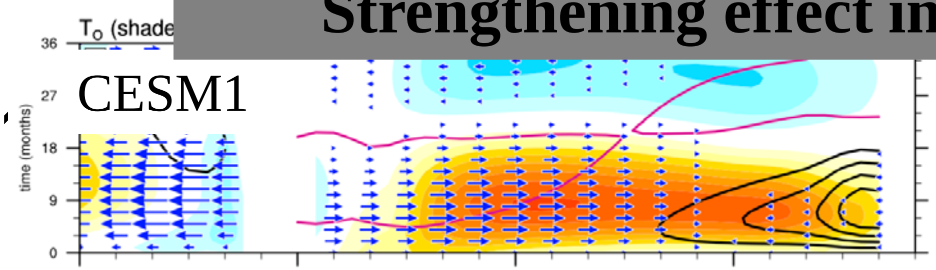
OBS



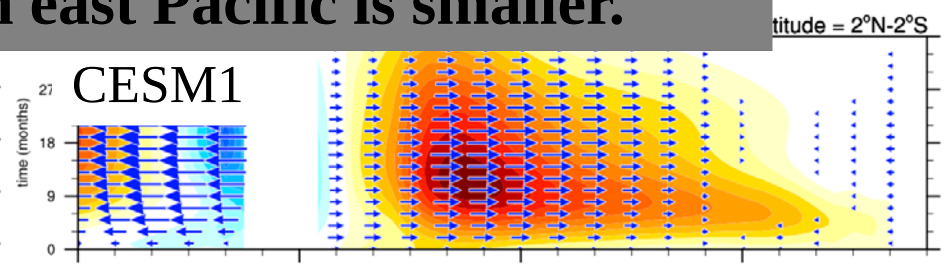


In these models (and some others), thermocline term (all effects) *weakens* ENSO more in central Pacific. Strengthening effect in east Pacific is smaller.

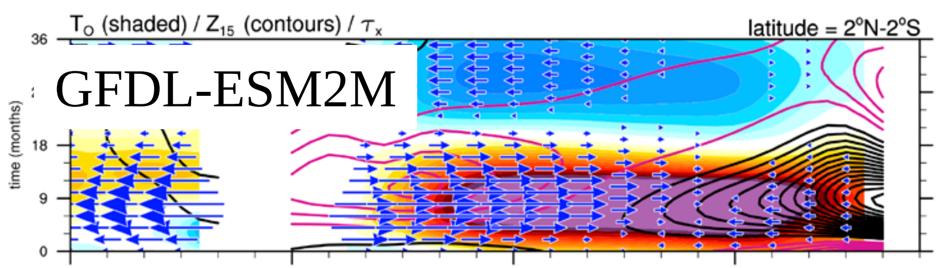
CESM1



CESM1



GFDL-ESM2M



T_O (shaded) / Z_{15} (contours) / τ_x latitude = 2°N-2°S

GFDL-ESM2M

Conclusion 2

- **Coupled GCMs may not represent observed dynamics sufficiently well enough to study ENSO diversity if they have:**
  - Strong and unrealistic Nino4 variability within “east Pacific” ENSO events that swamps more purely “central Pacific” ENSO variability
  - Too strong “nonlinear” relationship in PC1/2 plane
  - Optimal structure for central Pacific ENSO in west Pacific
  - Thermocline term acting *damps too much* in central Pacific and drives *growth too little* in east Pacific

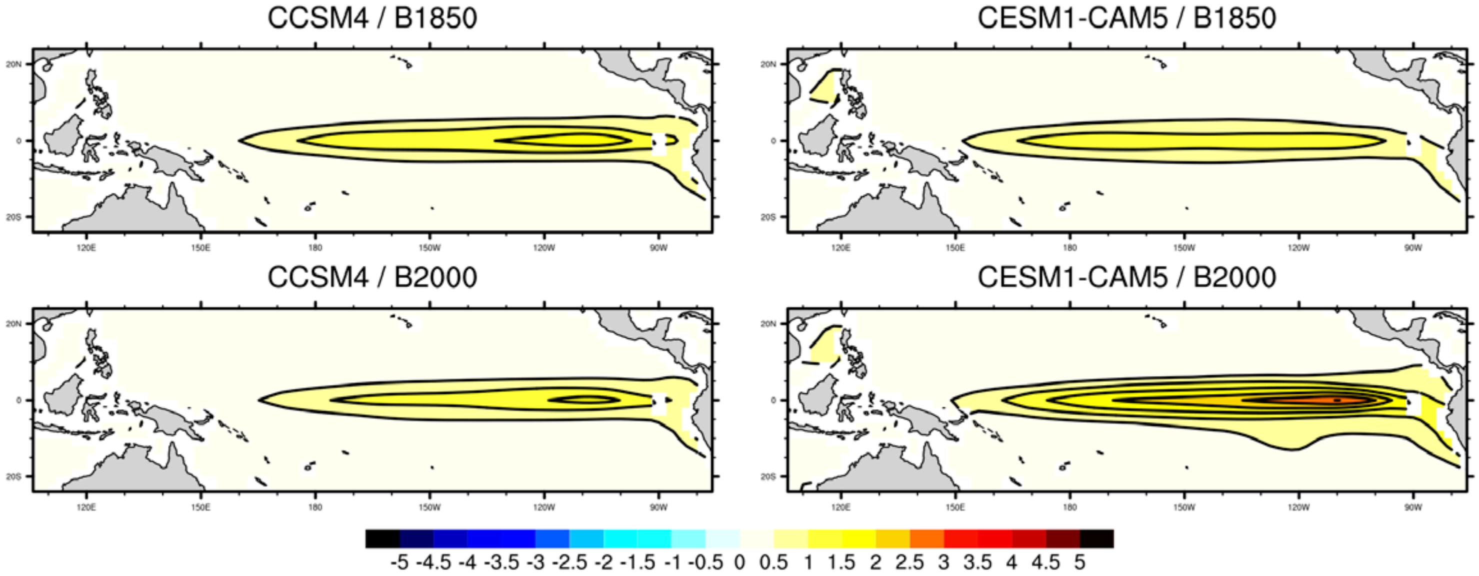
Compare LIMs constructed from:

- CCSM4, 1100 yr run under radiative conditions from
  - 1850 (B1850)
  - 2000 (B2000)
- CESM1-CAM5, 700 yr run under radiative conditions from
  - 1850 (B1850)
  - 2000 (B2000)

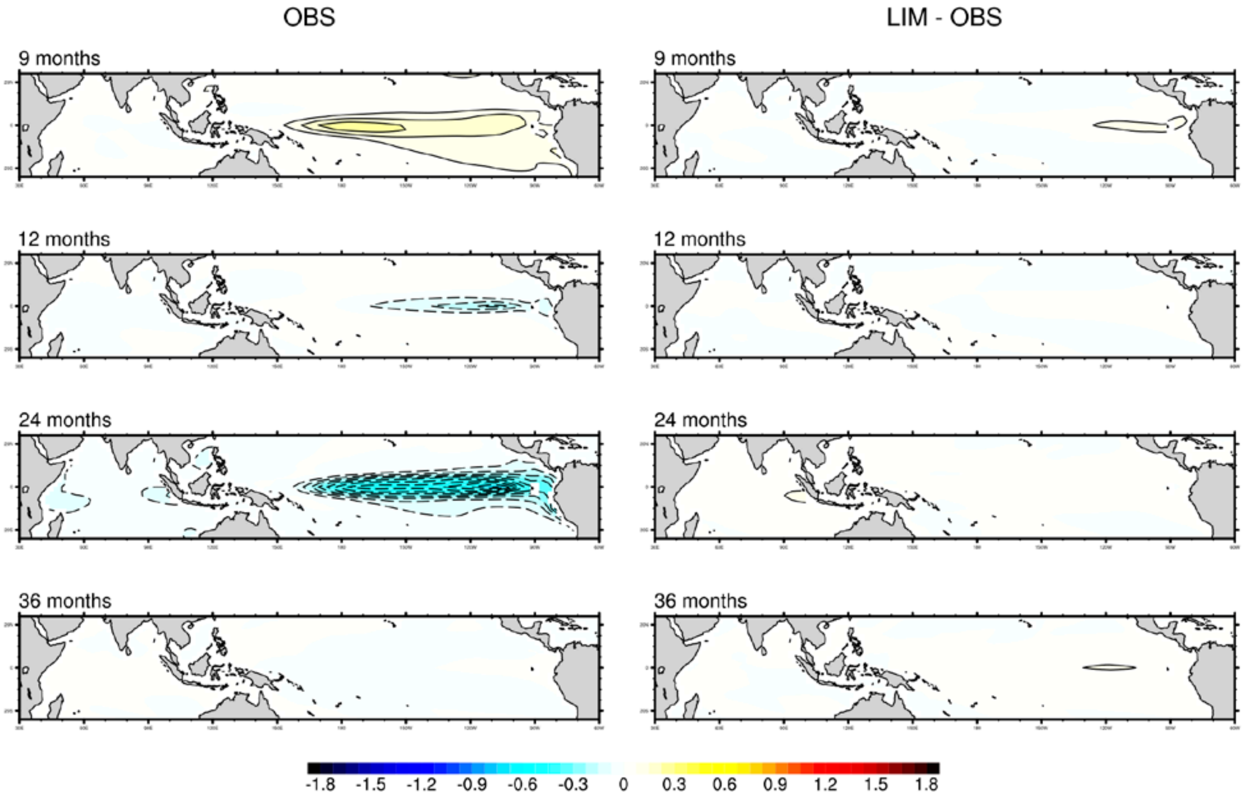
For all four cases, construct LIM from each half of record, use to make “forecasts” for the other half.

Variables are monthly anomalies of SST, Z20, and zonal wind stress.

Tropical SST variance

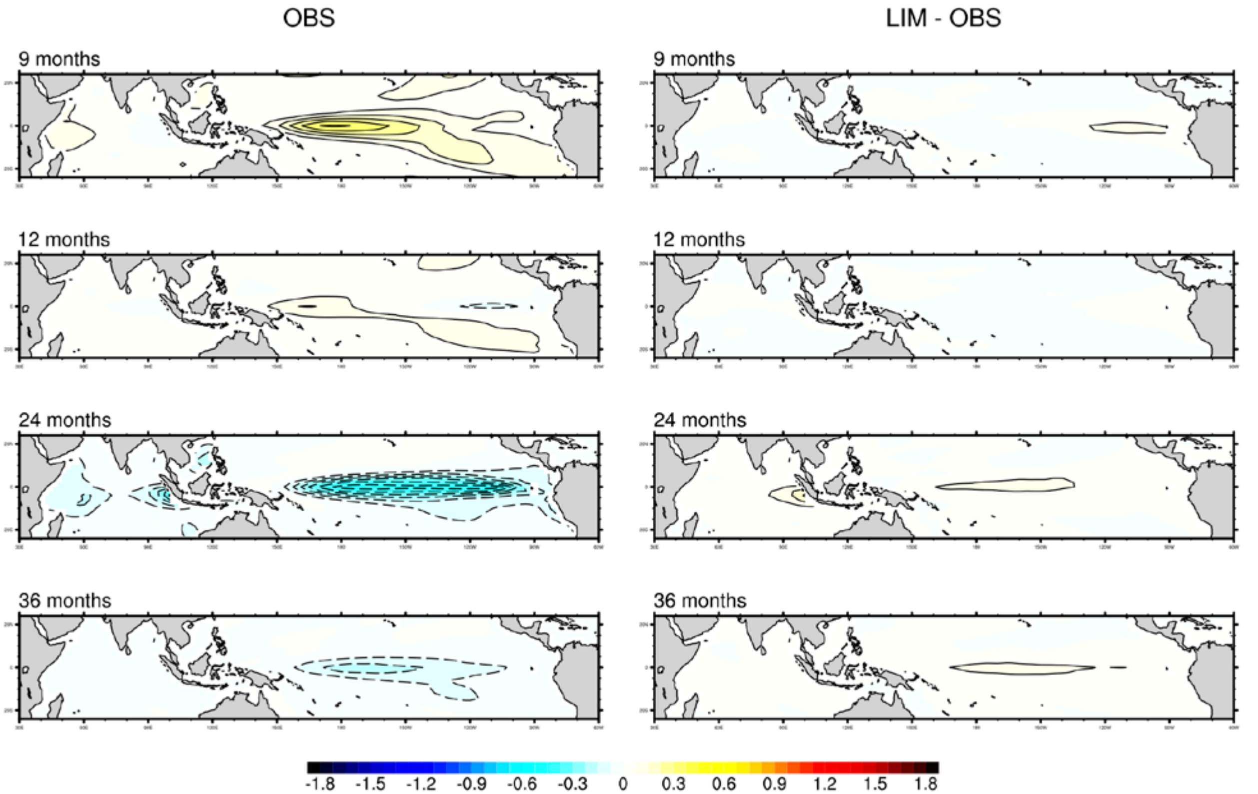


Test LIM: Tropical SST lag covariance



CCSM4 B1850

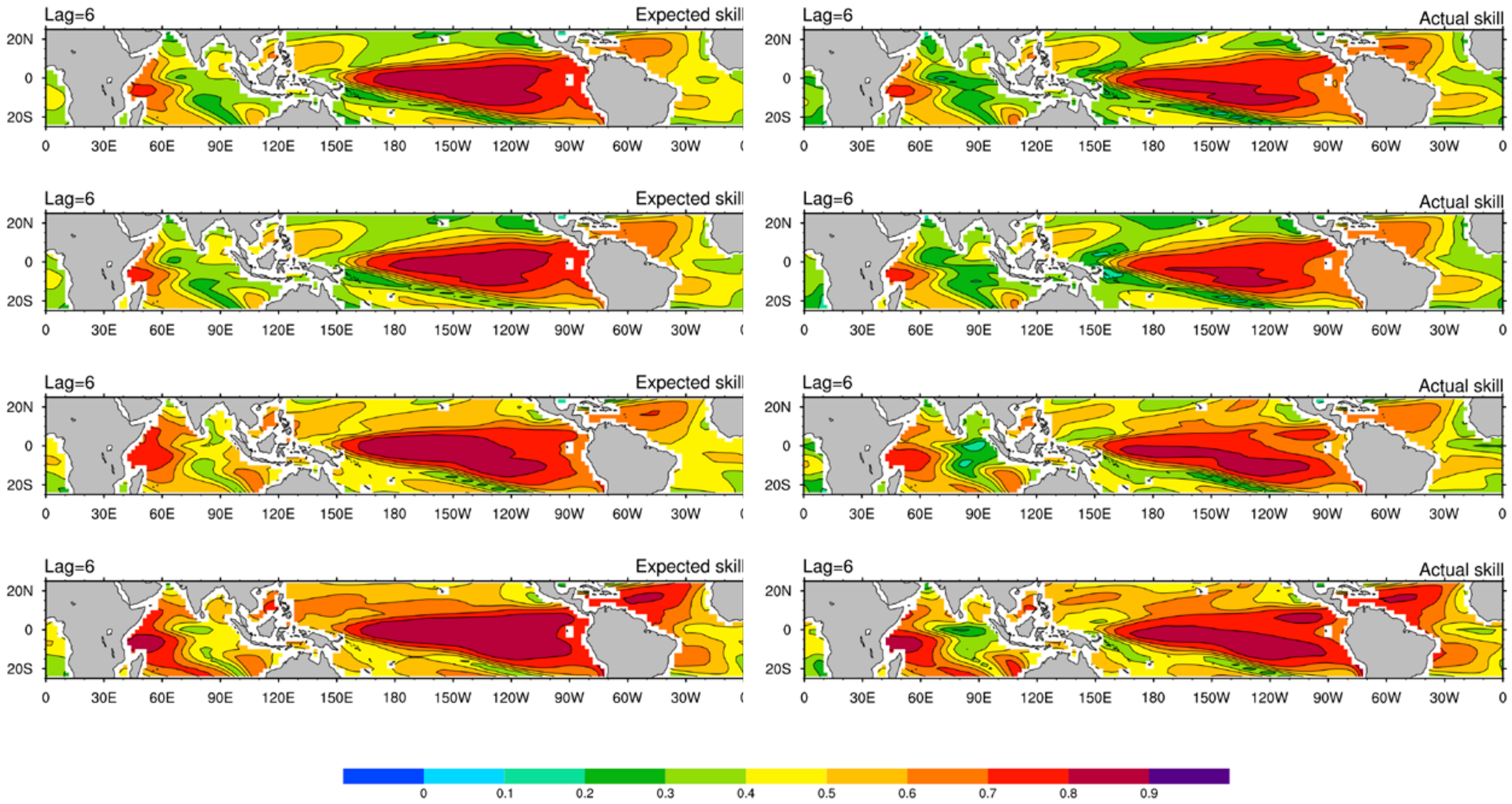
Test LIM: Tropical SST lag covariance



CESM1-CAM5 B1850

Comparing predicted and actual forecast skill, or: some places are more predictable than others

Predicted Month 6 LIM skill | Actual Month 6 LIM skill



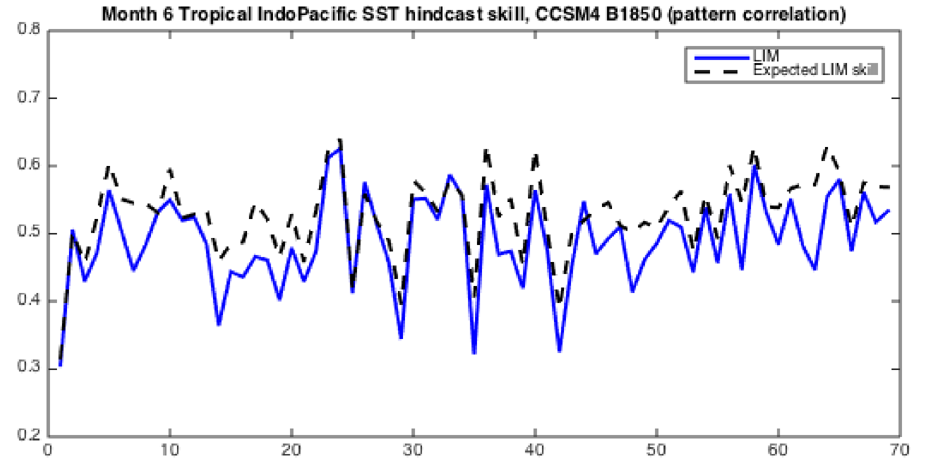
Comparing predicted and actual CCSM4 forecast skill

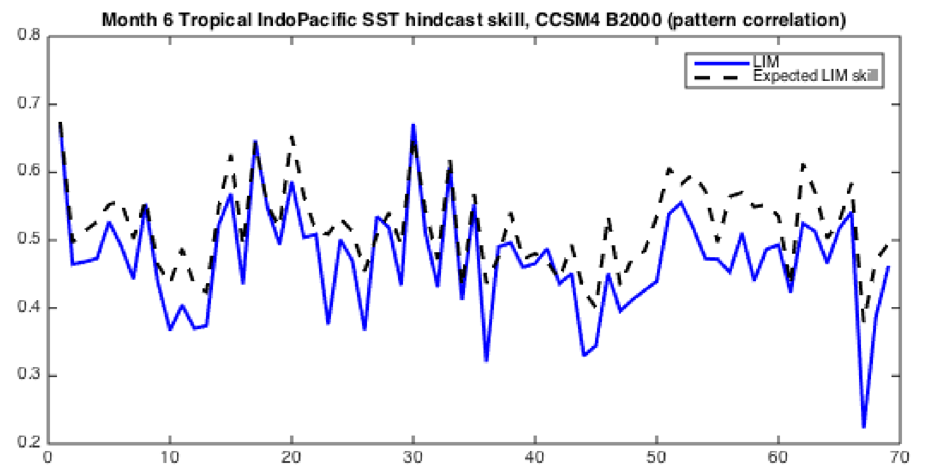
Skill averaged by **decade** (based on initialization date)

Actual skill: Pattern correlation of tropical SST anomaly with SST forecast anomaly

Blue: LIM (0.49/0.47)

Predicted: ρ_∞ (0.53/0.52)





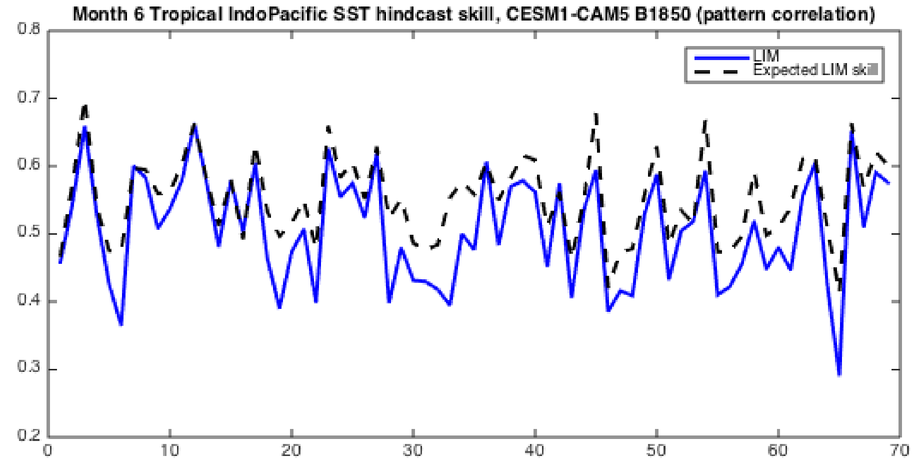
LIM skill correlated with ρ_∞ at 0.9

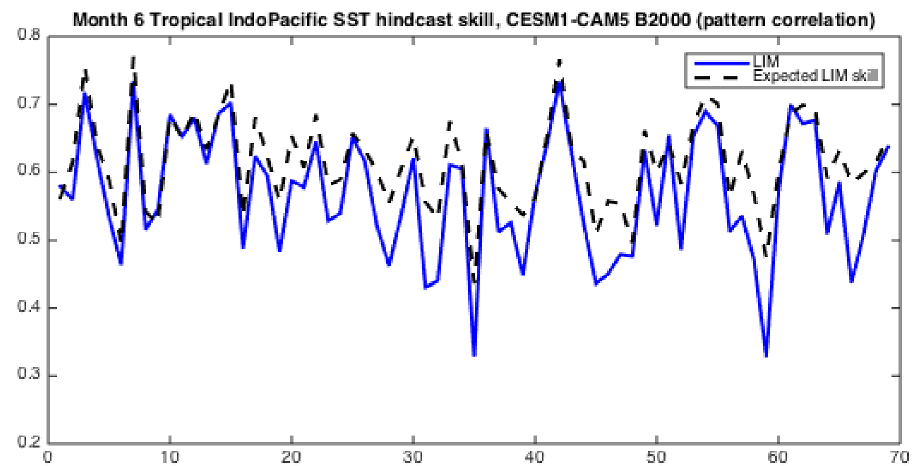
Comparing predicted and actual forecast skill, or: some years are more predictable than others

Skill averaged by decade (based on initialization date)

Actual skill: Pattern correlation of tropical SST anomaly with SST forecast anomaly

Blue: LIM (0.51/0.57)
Predicted: ρ_∞ (0.55/0.62)





LIM skill correlated with ρ_∞ at 0.9

Conclusions

- LIM useful for diagnosis of predictability, ***because*** its forecast skill is comparable with coupled GCMs and it reproduces observed spatio-temporal statistics
- To use coupled CGMs to investigate changes in ENSO and its predictability we first need to gauge how well do they reproduce observed ENSO dynamics
- Year-to-year variations in forecast skill in the last few decades may be due to random variations in initial conditions and not necessarily to long-term “base state” changes
- Even if “base state” changes drive variations in forecast skill, these may be swamped by random variations in initial conditions