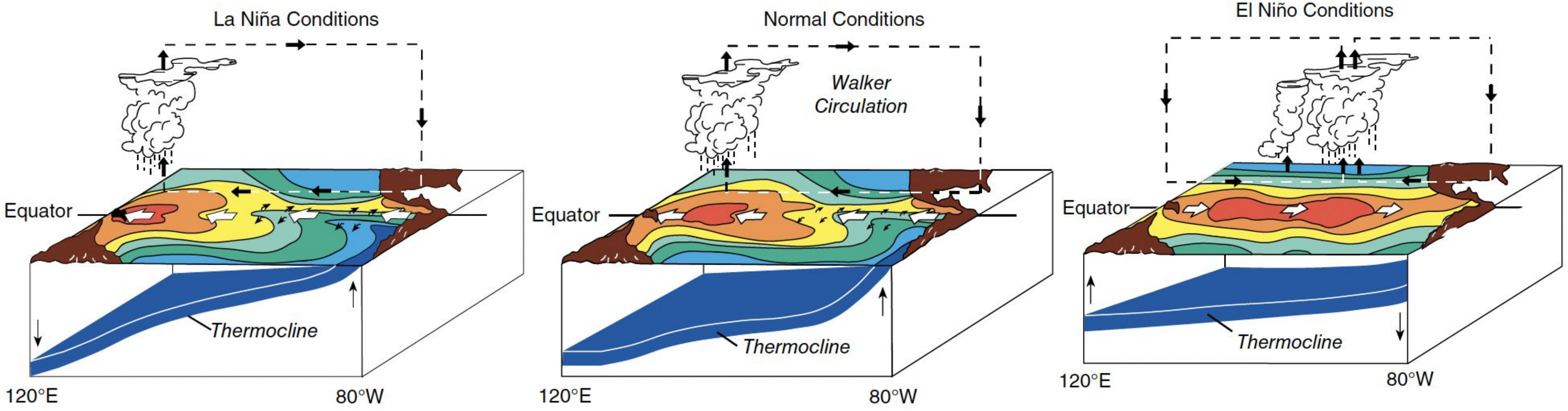
Fast-track evaluation of ENSO performance using a comprehensive suite of GFDL models

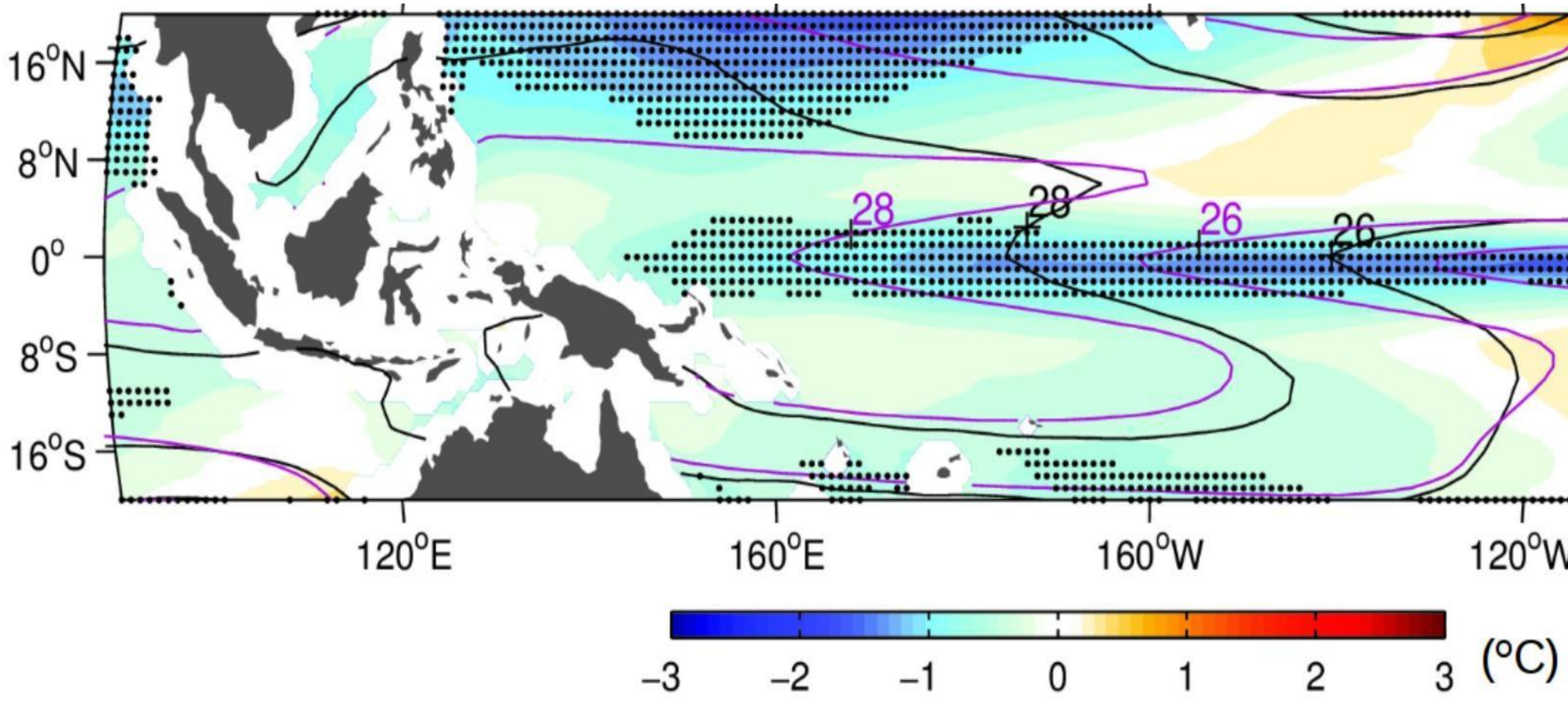
Jiale Lou [1] , and Andrew Wittenberg [2]

[1]Atmospheric and Oceanic Sciences Program, Princeton University, Princeton, NJ, USA
[2] NOAA Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA



Schematic adapted from McPhaden et al. (2020)
"Introduction to El Niño Southern Oscillation in a Changing Climate"

Introduction: Common tropical Pacific biases in CGCMs



Santoso et al. (BAMS 2019)

20-model mean for CMIP5 historical runs, relative to ERSST.v5. SST contours are for obs (black) and models (purple).
Stippling: at least 90% of models have bias of same sign.

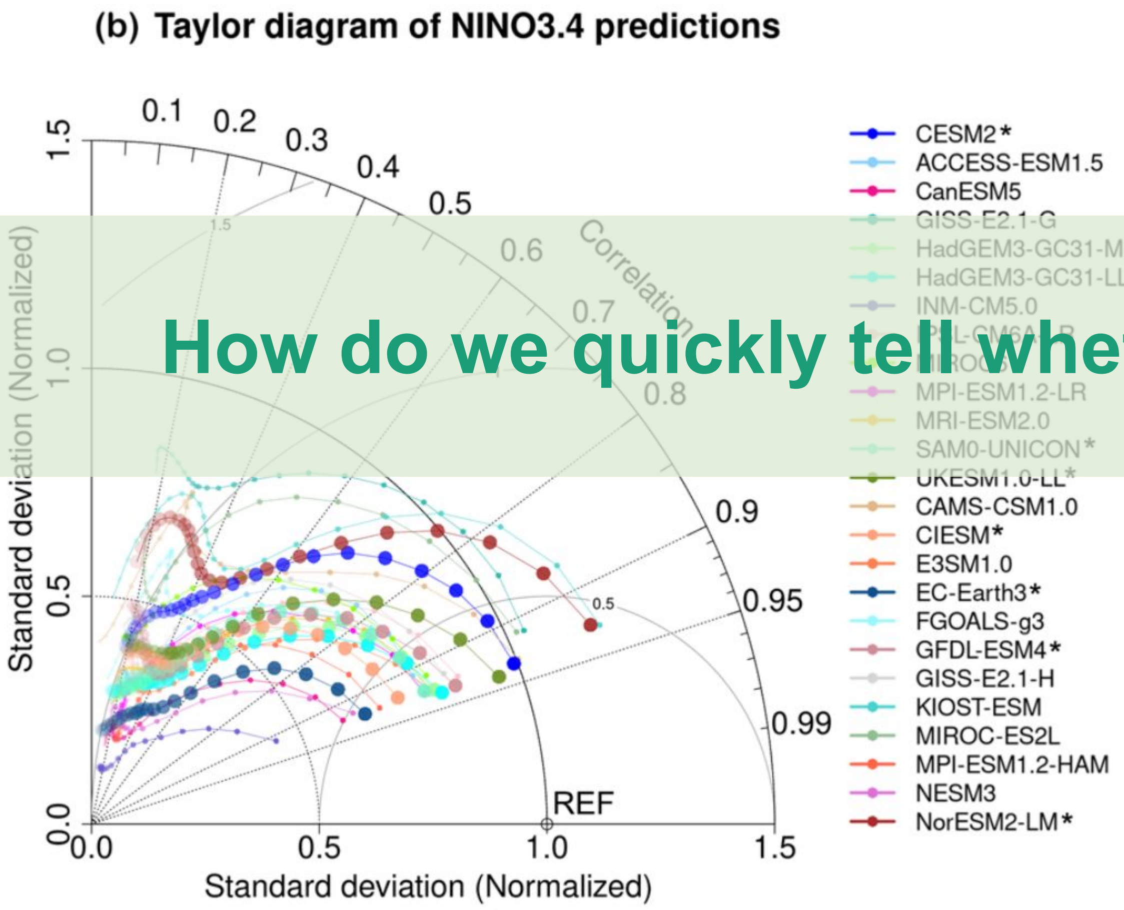
Equatorial cold SST bias, eastern warm SST bias, excessive trade winds, double ITCZ ...

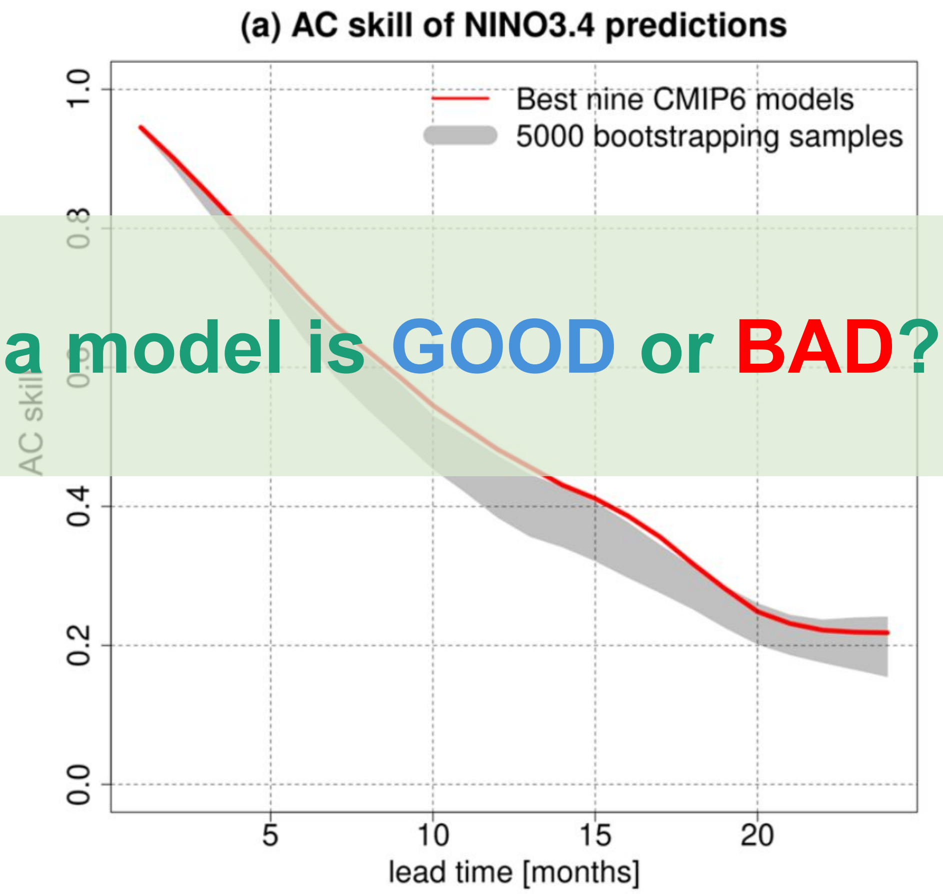
→ Weak ENSO coupling, damping, diversity, asymmetry; shifted teleconnections
→ Affects data assimilation & initialization, induces drifts & shocks in forecasts

Slide adapted from Andrew Wittenberg (GFDL)

Data-driven ENSO forecast skill depends on model performance

Data-driven ENSO forecast skill is fundamentally constrained by the quality of ENSO simulation in climate models. Methods such as model-analog forecasting, machine learning, and statistical predictions rely on the model's ability to realistically represent ENSO mean state, variability, and teleconnection structures.





How do we quickly tell whether a model is GOOD or BAD?

Lou, Newman, and Hoell (2023; *npj clim. atmos. sci.*)

How do we quickly tell whether a model is GOOD or BAD?

BAMS Article

Evaluating Climate Models with the CLIVAR 2020 ENSO Metrics Package

Yann Y. Planton, Eric Guilyardi, Andrew T. Wittenberg, Jiwoo Lee, Peter J. Gleckler, Tobias Bayr, Shayne McGregor, Michael J. McPhaden, Scott Power, Romain Roehrig, Jérôme Vialard, and Aurore Voldoire

$$Standardized\_scores = \frac{metric - multi\_model\_mean}{\sigma}$$

Red = worse
Blue = better
(closer to obs)

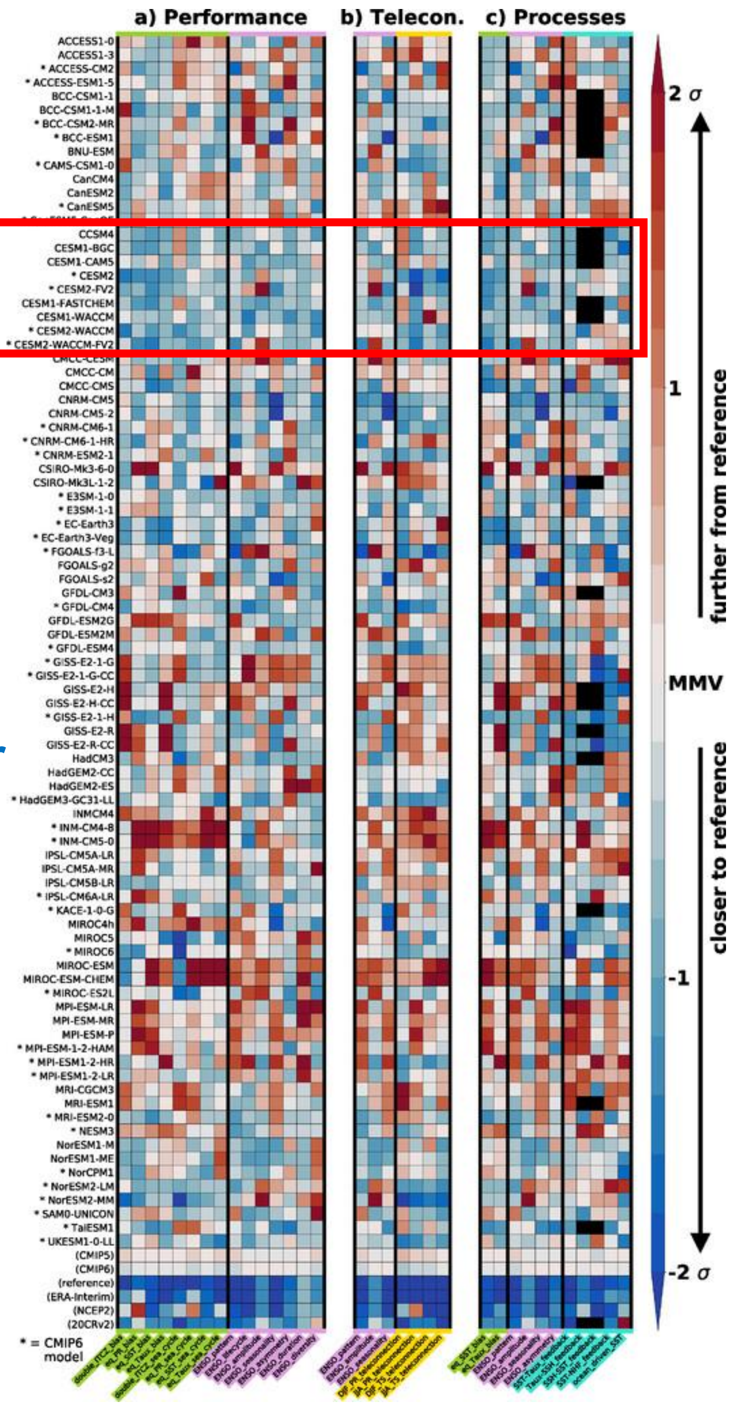
metric: either **RMSE** or **Absolute Percentage Error (APE)** computed between model and observations.

$$RMSE = \sqrt{<(model - obs)^2>}$$

Such as mean state biases, seasonal cycle, ENSO SSTa pattern, ENSO lifecycle.

$$APE = \frac{|model - obs|}{|obs|} \times 100\%$$

Such as ENSO skewness, amplitude, seasonality, asymmetry, duration, diversity, and ENSO feedback processes.



Newly developed GFDL CGCMs

- Large ensembles
- Resolution

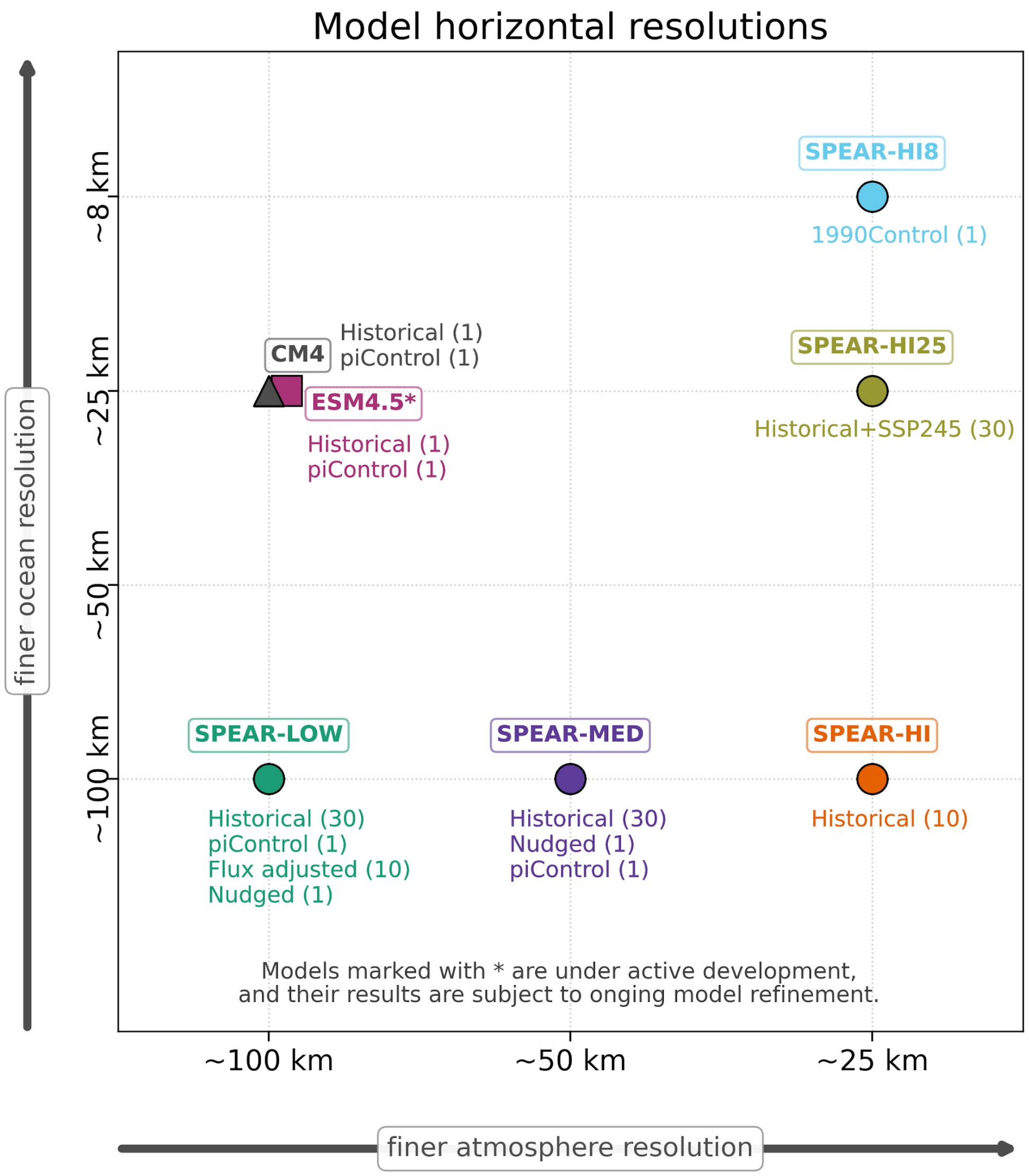
Comprehensiveness

- Physics innovation
- Resolution

	SHiELD (2020 & onward) Weather to Seasonal Data-Initialized Physical Prediction	**SPEAR** (2020 & onward) Seasonal to Multi-decadal Data-Initialized Physical Prediction	**ESM4.5** (2025) Decadal to Century Full Earth System Projection	**CM5** (2026, 2028) Decadal to Century Physical Climate Sea Level
FV3 dycore Atmosphere	**SHiELD** 3 to 13 km; 91 Level	**AM4** 25 to 100 km; 33 Level	**AM4.5** 100 km; 49 Level	**AM5** 25 or 100 km; 65 Level
Atmospheric Chemistry	Simple Aerosols	Simple Chemistry & Aerosols	Full Chemistry & Aerosols	Simple Chemistry & Aerosols
LM4 Land	**NOAH LSM** (Initialized **LM4.2i** planned)	**LM4.0/LM4.2** Ecosystems	**LM4.5** Ecosystems, Fire, Snow	**LM4+** Orography Aware
MOM6 / SIS2 Ocean / sea-ice	**Mixed Layer** (**OM5** planned)	**OM4**-derived 1° to $^{1}/_{12}$°; 75 Layer	**OM5** ¼°; 75 Layer	**OM5** (non-Boussinesq) ¼° to $^{1}/_{12}$°; 75 Layer
FMS Coupler & Infrastructure	**Atmospheric Ensemble Data Assimilation**	**Ensemble Data Assimilation**	**COBALTv3 Ocean Ecosystems**	**Interactive Dynamic Ice Sheets**

Slide adapted from GFDL 5-year Review (2025)

GFDL CGCM Overview



SPEAR: Designed for **predictions** (e.g., NMME) and **projections** (large ensembles) of climate risks at **seasonal to centennial** scales.

AM4 atmosphere (33 levels), **LM4** land**, MOM6** ocean (75 hybrid levels), **SIS** sea ice

Physical components similar to CM4 (Held et al. 2019) & ESM4.1 (Dunne et al. 2020); but different resolution

		Atmos grid	Ocean grid (and equatorial Δy)
SPEAR	**SPEAR_LO**	100 km	100 (33) km
SPEAR	**SPEAR_MED**	50 km	100 (33) km
SPEAR	**SPEAR_HI**	25 km	100 (33) km
SPEAR	**SPEAR_HI25**	25 km	25 km
SPEAR	**SPEAR_HI8**	25 km	8 km
	CM4	100 km	25 km
	ESM4.5*	100 km	25 km

SPEAR: Seamless System for **P**rediction and **EA**rth System **R**esearch (Delworth et al. 2020)
CM4: CMIP6 archive (Held et al. 2019)
ESM4.5*: Earth System Model 4.5 (under active development)

Slide adapted from Tom Delworth (GFDL) and Andrew Wittenberg (GFDL)

Observational references

What we compare models against in the ENSO diagnostics

SST:
ERA5
HadISST
ERSSTv5

SSH:
GODAS

Precipitation:
ERA5
GPCPv2.3
GPCPv3.2

Wind stress:
ERA5

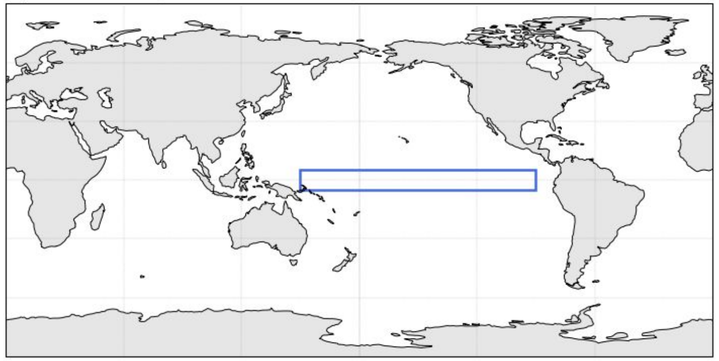
Radiative fluxes (SW/LW):
ERA5

Heat fluxes (Latent and Sensible):
ERA5

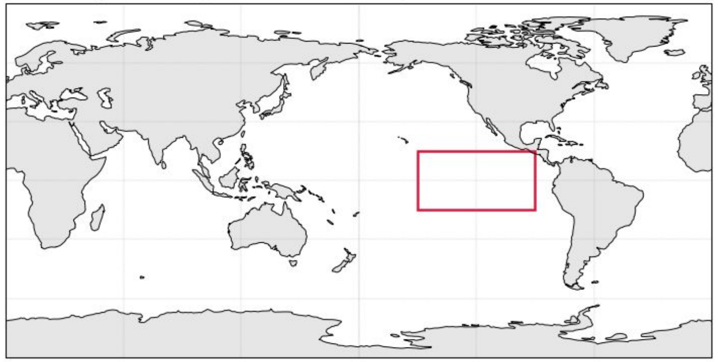
Monthly data; Remapped to 1°x1° gridding; SST period: 1940-2014; Precip period: 1980-2014

Housekeeping:

Equatorial Pacific [5°S–5°N, 150°E–90°W]

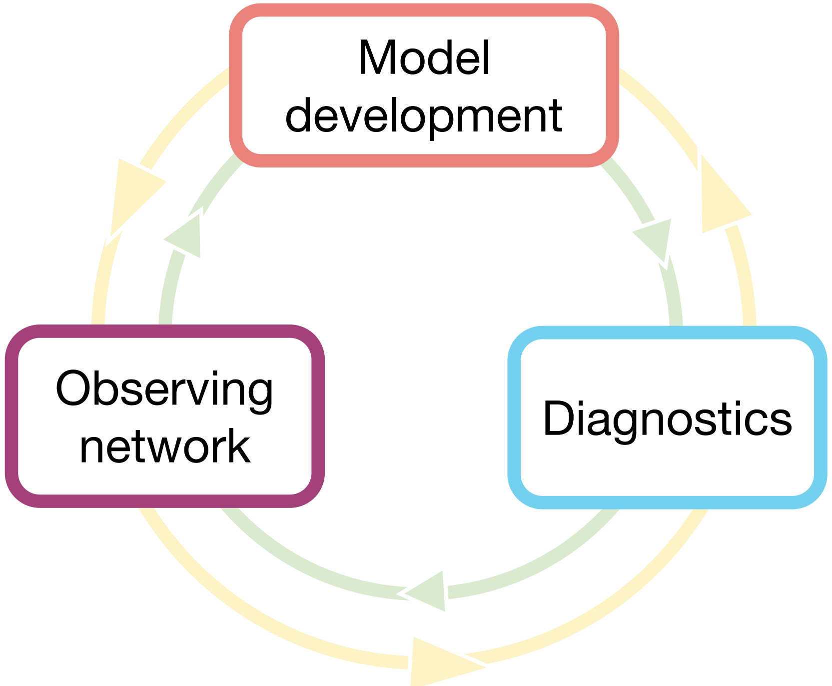


Eastern tropical Pacific [15°S–15°N, 150–90°W]

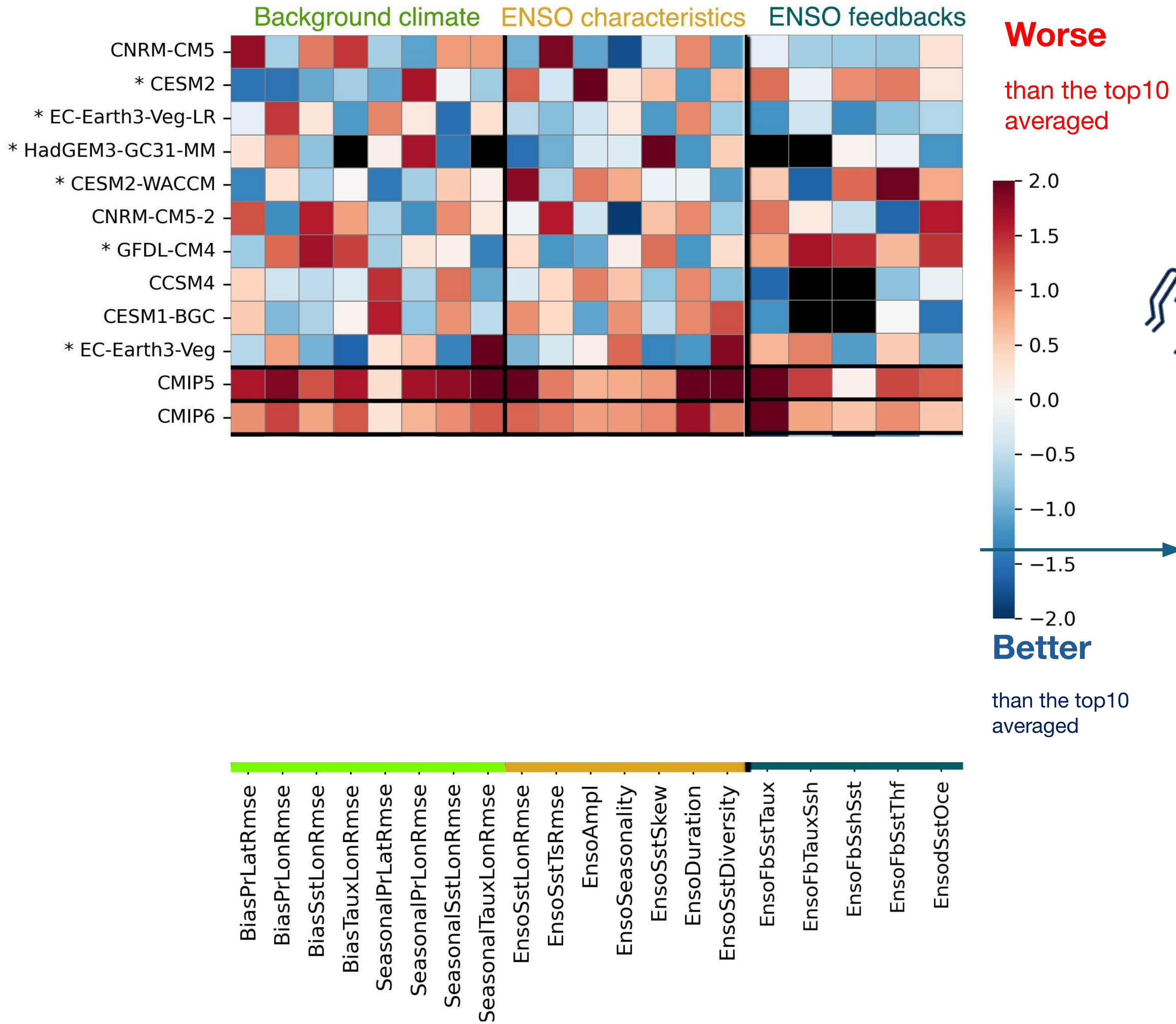


Research questions

- Does increasing model resolution systematically improve ENSO simulation?
- How robust are ENSO simulations across large ensembles?
- How can diagnosed model biases inform observing-network design in the tropical Pacific such as Tropical Pacific Observing System (TPOS) Equatorial Pacific Experiment (TEPEX)?



Results: ENSO metric scores for top-10 CMIP and GFDL models



Top-10 CMIP models are selected by ranking the row-wise average of the metric scores.

NCAR's models in top-10: *CESM2, *CESM2-WACCM, CCSM4, CESM1-BGC

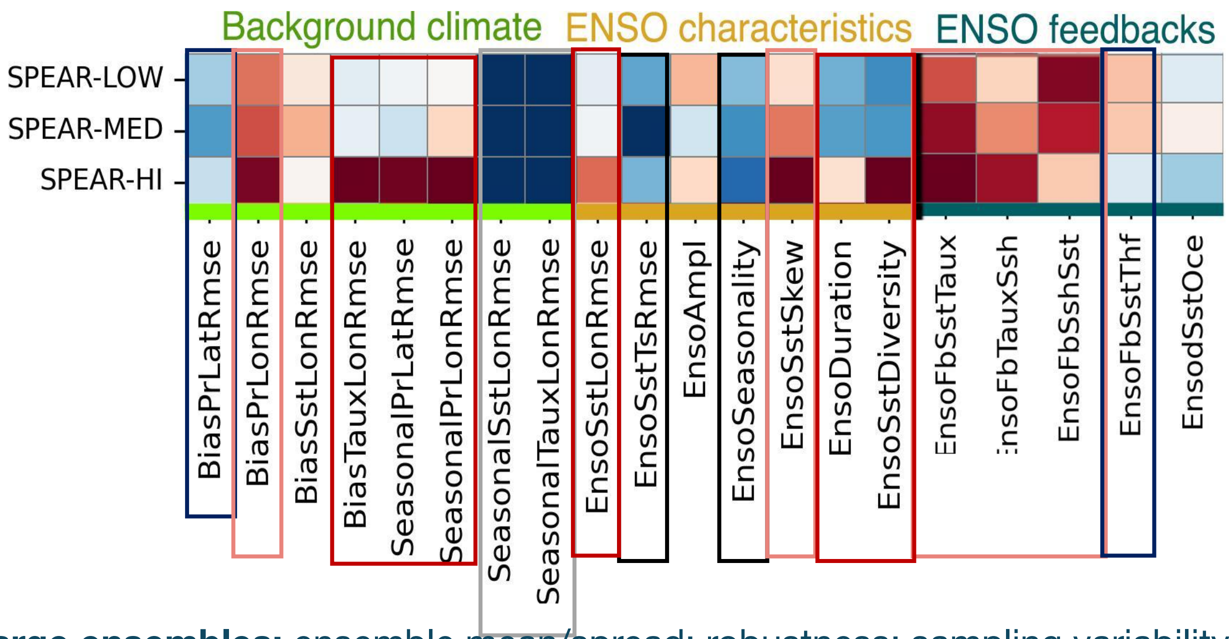
Nudged runs
SST, SSS, and wind stress are nudged to the observations

Flux-adjusted
Remove climatological SST/SSS/stress biases
Xian Wu et al., (submitted)

Free coupled runs
With increased horizontal resolutions

Large ensembles: ensemble mean/spread; robustness; sampling variability

Atmosphere resolution: SPEAR-LOW, -MED, -HI (atmos: 100km□50km□25km)



Large ensembles: ensemble mean/spread; robustness; sampling variability

Consistently good (better than top-10 averaged):

Climatological Pr in the eastern tropical Pacific (15S-15N)
Seasonal cycle of SST and Taux along the eqPac
ENSO lifecycle
ENSO seasonality

Consistently bad (worse than top-10 averaged):

Climatological Pr along the eqPac (bad and gets worse)
ENSO skewness (bad and gets worse)
SST-taux feedback
ENSO taux-SSH feedback
ENSO SSH-SST feedback

Getting better with finer atmos resolution:

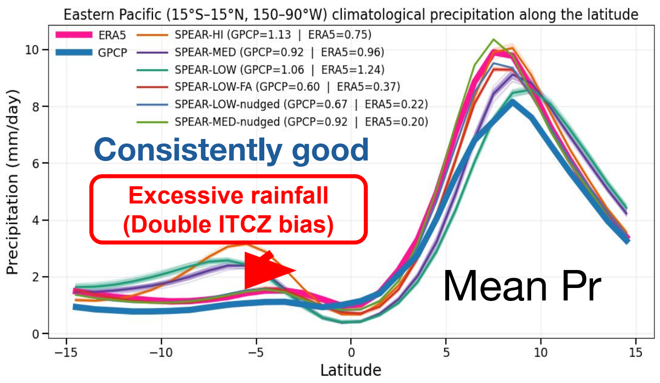
Thermal damping feedback (SST-Qnet feedback)

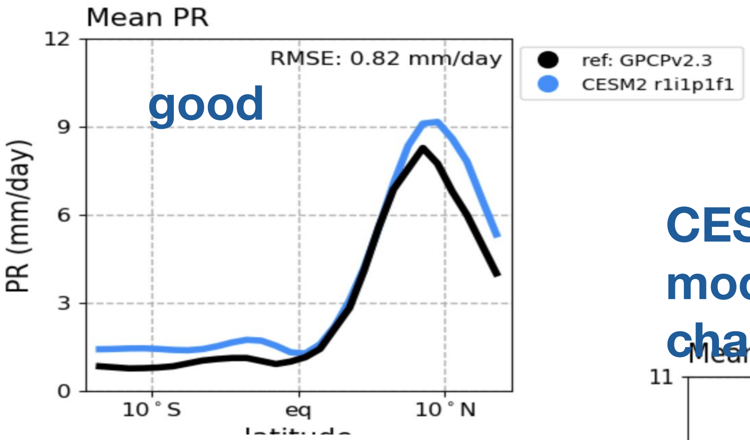
Finer atmosphere ≠ better ENSO
BUT WHY?

Getting worse with finer atmos resolution:

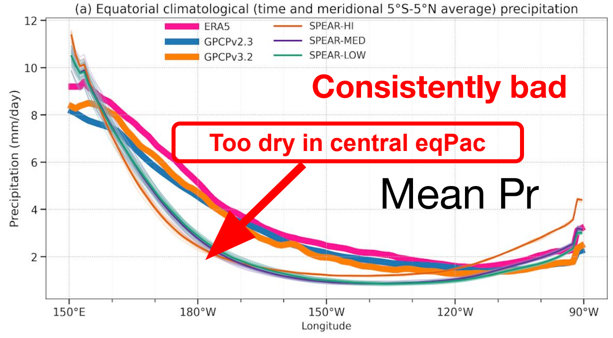
Climatological Taux along the eqPac
Seasonal cycle of Pr in the eastern tropical Pacific (15S- 15N) and along the eqPac
ENSO pattern
ENSO duration
ENSO diversity

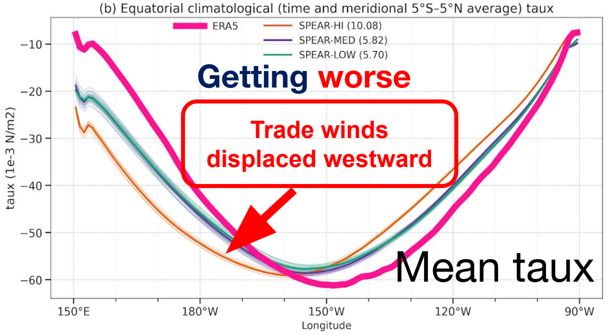
Dive-down diagnostics (Mean state)

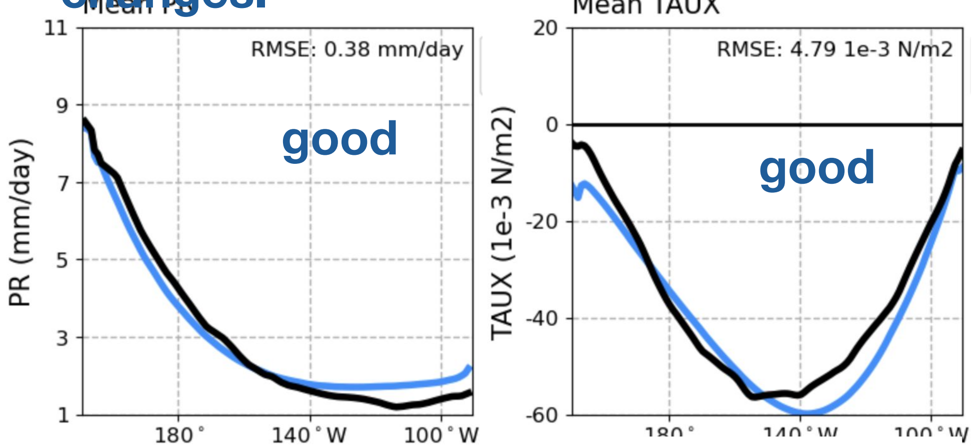




CESM2 is among the best-performing models for tropical mean-state changes.







Rainfall biases persist even when SST is nudged toward observations (top panel)

Linked structural biases:
Overly strong trade winds over the western Pacific → westward-shifted convection → overly sharp rainfall gradients.

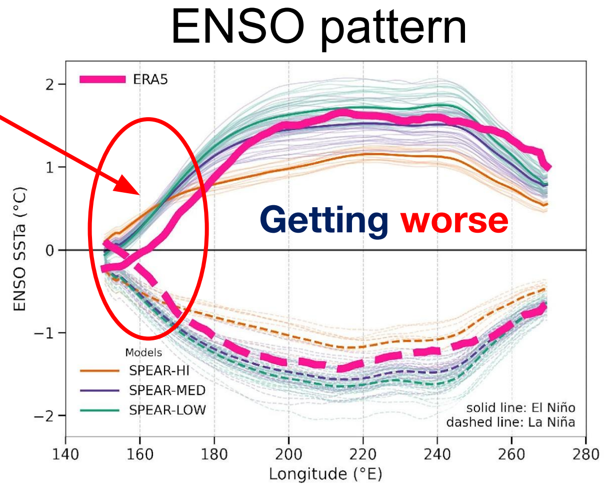
Excessive seasonal variation of ITCZs

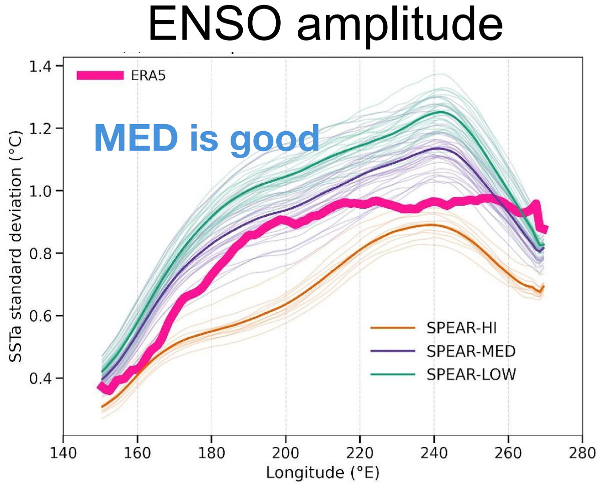
Too strong in the east and west eqPac

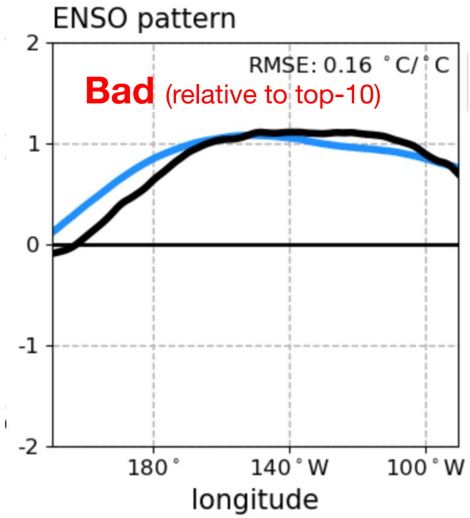
Dive-down diagnostics (ENSO characteristics)

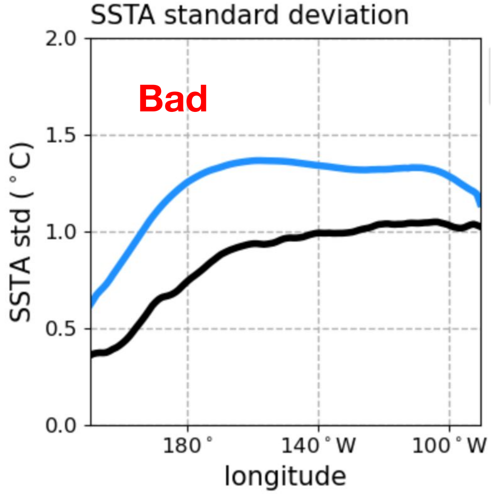
Westward extension

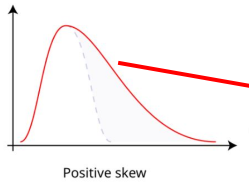
Too weak in SPEAR-HI

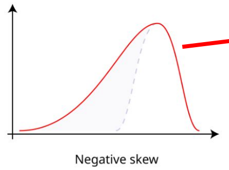




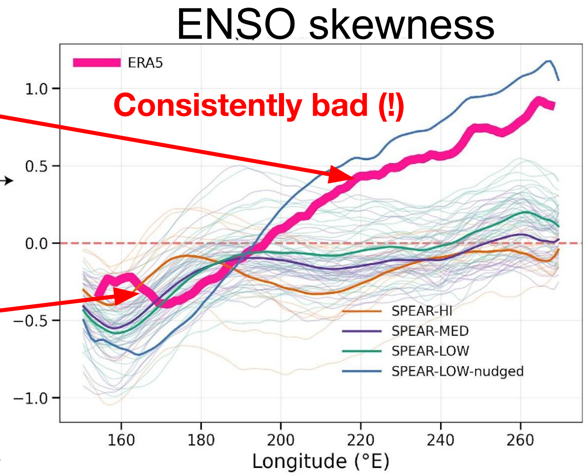


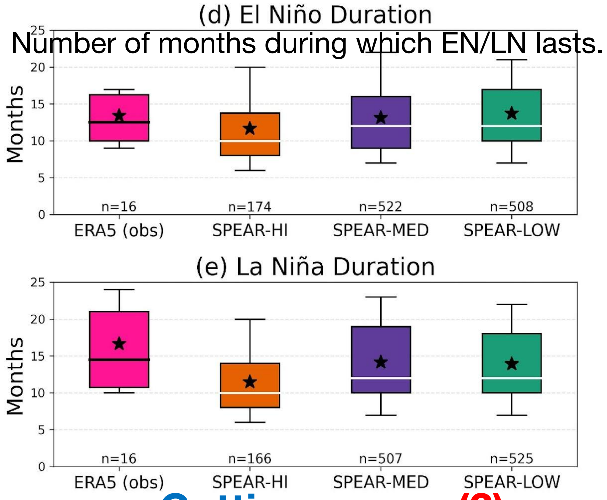




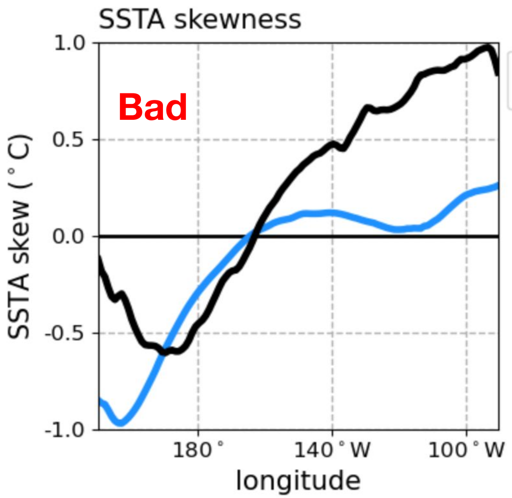


For illustration only
not the actual PDF



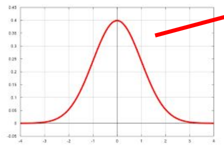


Getting worse (?)



CESM2:

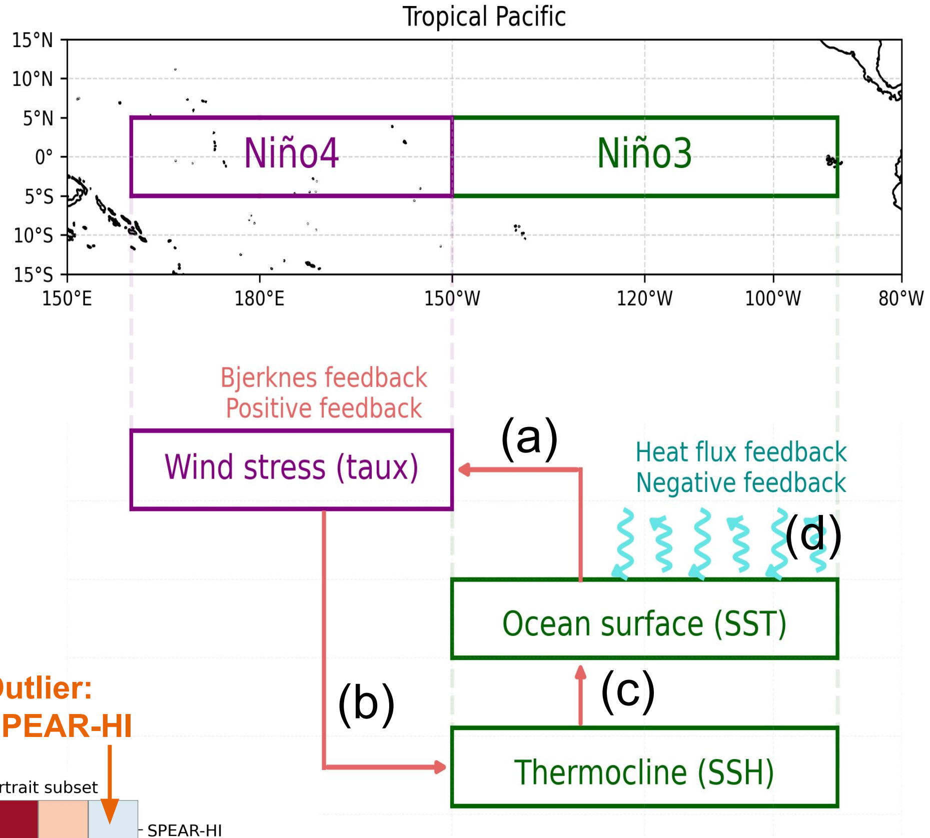
- ENSO westward extension
- Too strong ENSO amplitude
- Too symmetric over central eastern Pacific.
- Good ENSO duration (not shown).

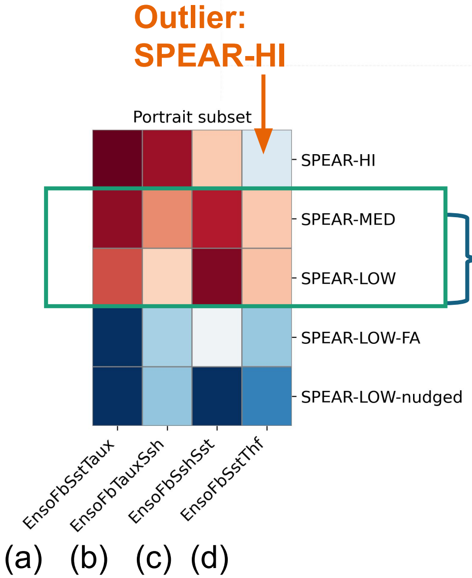


Too symmetric
Wide spread;
No ensemble members fall within the observed range

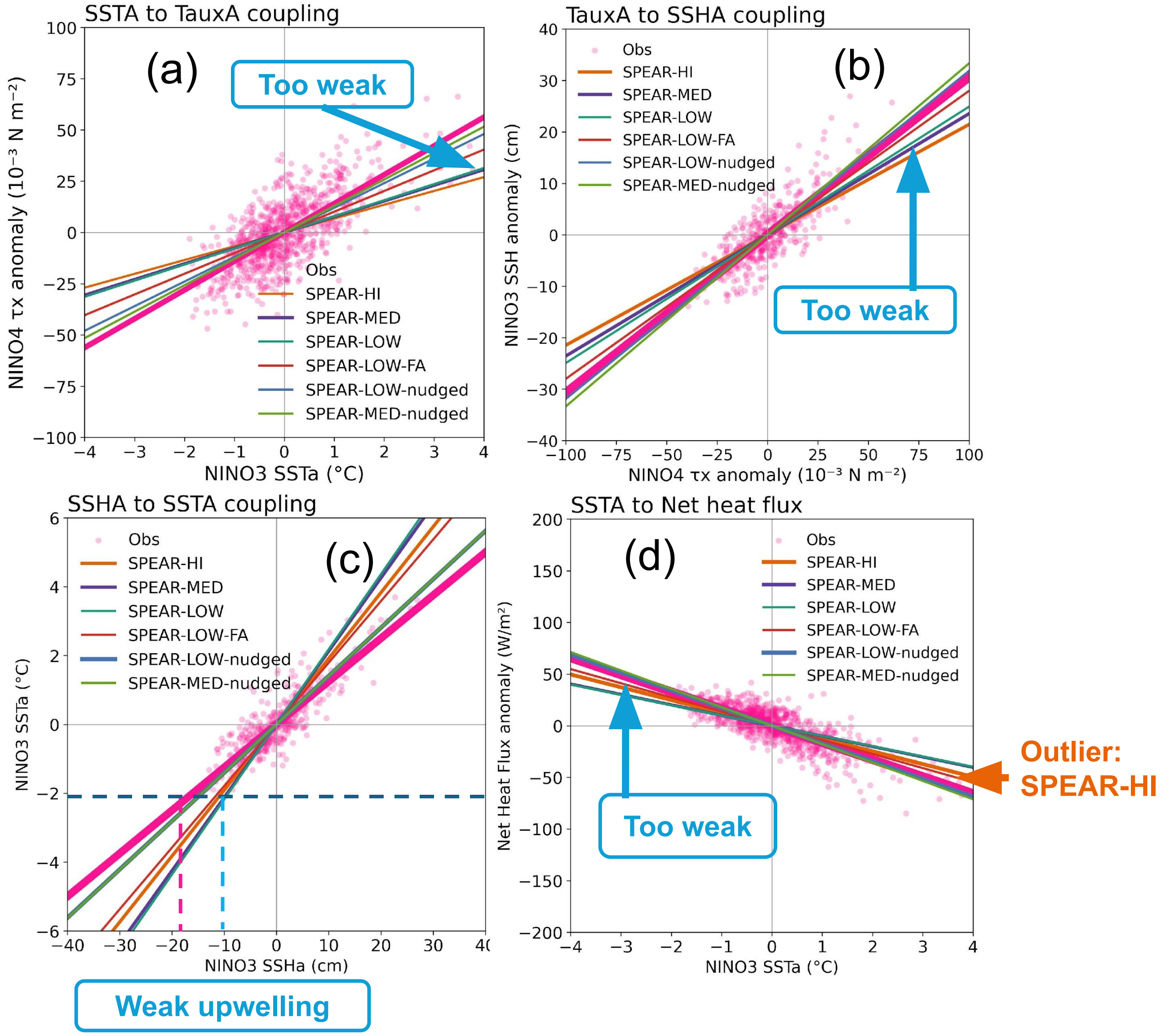
Despite the shorter ENSO duration in SPEAR-HI, this likely reflects sampling uncertainty, as the model spread overlaps with the observed range.

Dive-down diagnostics (ENSO feedback processes)





Error compensation: an overly weak positive feedback (Bjerknes loop) offsets an overly weak negative feedback (thermal damping), yielding seemingly realistic ENSO variability.



WHY increasing atmos resolution does not improve ENSO simulations?

Resolution matters but parameterization schemes matter more

At finer grid spacing (SPEAR-HI vs. -MED and -LOW): The resolved flow captures more realistic small-scale variability. The sub-grid diffusion and microphysics settings can become too weak or too active, leading to noisy or unrealistic behavior.

Past tunings:

Sub-grid diffusion parameterization: Tuned up □ Stabilized large-scale precipitation

Droplet radius: Tuned up □ Harder for clouds to rain

Fallout speed: Tuned up □ accelerates condensate removal once rain forms

Delayed precip onset

Intensified precip events

✓ **Strength:** Extreme precipitation representation improves *(*Hiroyuki Murakami; Bor-Ting Joh*)*

✗ **Weaknesses:** Seasonal cycle of precipitation too variable; ENSO variability too weak

Excessive rainfall □ |∂Q′/∂T′| increases □ Enhanced thermal damping □ Weak ENSO variability

(Negative feedback)

Summary

The ENSO diagnostics package (Planton et al., 2020, BAMS) enables rapid, fit-for-purpose selection and comparison of models.

SPEAR performs strongly for ENSO SST metrics, but has rainfall biases across resolutions, even when SST is nudged toward observations. This suggests a need for further advances in atmospheric convection & cloud parameterizations, especially at high resolution.

Biases in simulated background climate contribute to biases in ENSO patterns and skewness. Error compensation among weak ENSO feedbacks can produce realistic ENSO amplitudes for the wrong reasons.

These results highlight a need for targeted observations & process studies of the air-sea transition zone in the equatorial Pacific – e.g., surface fluxes, convection & clouds, ocean mixing □ TPOS/TEPEX!

Next: Extend the ENSO diagnostics package to E3SM, focusing on cloud–radiation feedbacks as I join the University of Wyoming next month; How the diagnosed ENSO biases affect ENSO forecast skill.

Thank you!

CMIP5/6 ENSO Metrics summary
https://pcmdi.llnl.gov/research/metrics/enso
- **Interactive** dive-down diagnostics
https://pcmdi.llnl.gov/pmp-preliminary-results/interactive_plot/portrait_plot/enso_metric/enso_metrics_interactive_portrait_plots_v20231121.html
- **Pre-computed CMIP-archived model metrics for download** (Excel)

Wiki: https://github.com/CLIVAR-PRP/ENSO_metrics/wiki
Software: https://github.com/CLIVAR-PRP/ENSO_metrics

Jiale Lou: jiale.lou@Princeton.edu
jiale.lou@noaa.gov
Website: https://www.jialelou.com

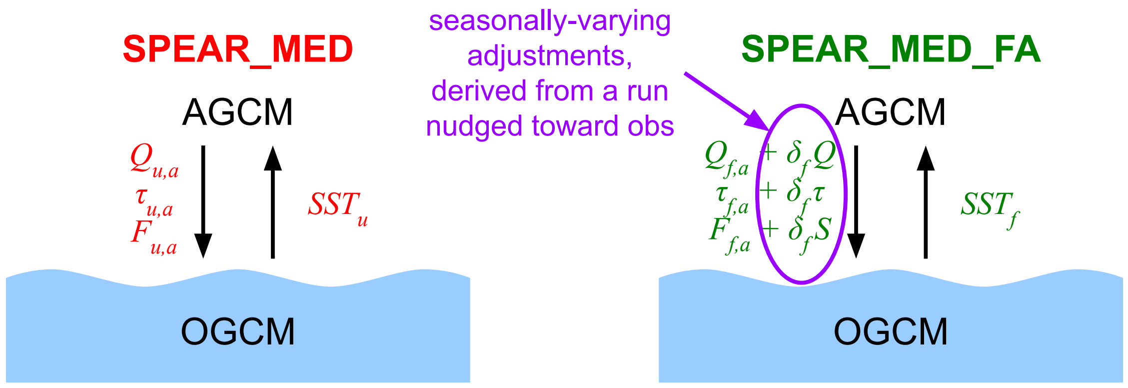
Supplementary materials
&
Reference sheet

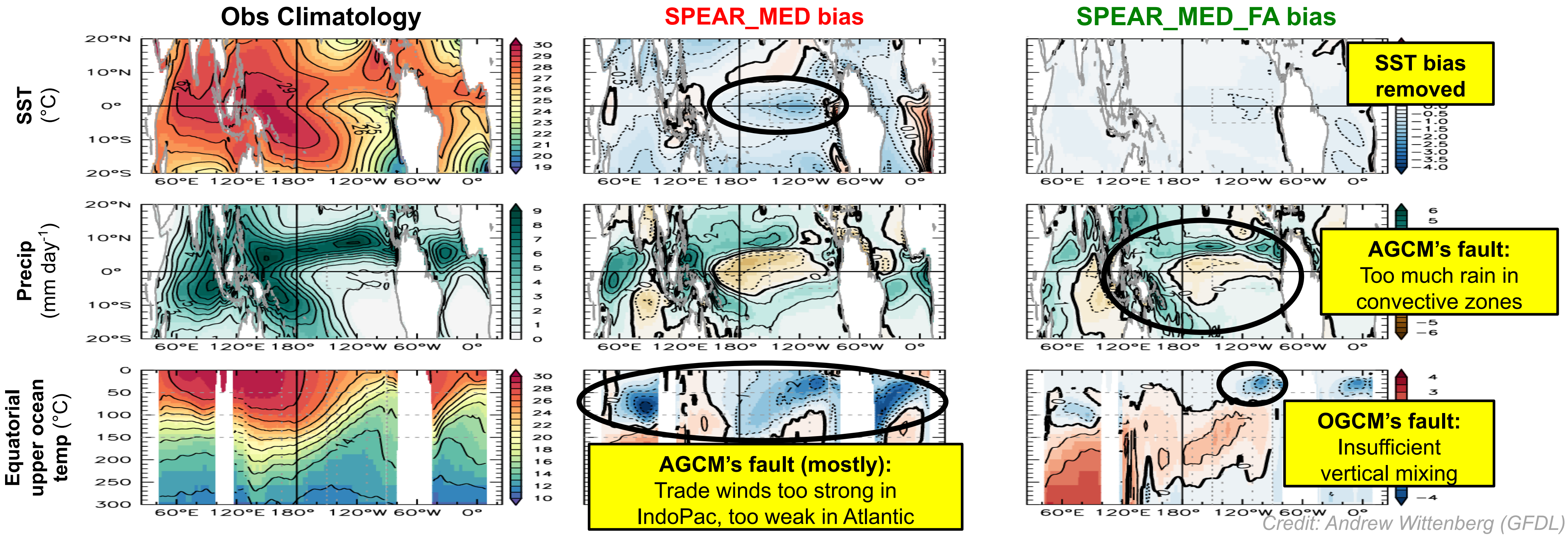
Flux-adjusted (FA) version of SPEAR

A. Wittenberg, F. Zeng, X. Wu, T. Delworth, W. Cooke

Adjust the surface heat, salt, & momentum fluxes passed from AGCM to OGCM
→ **corrects the climatologies of *tropical* SST/SSS and *global* wind stress.**

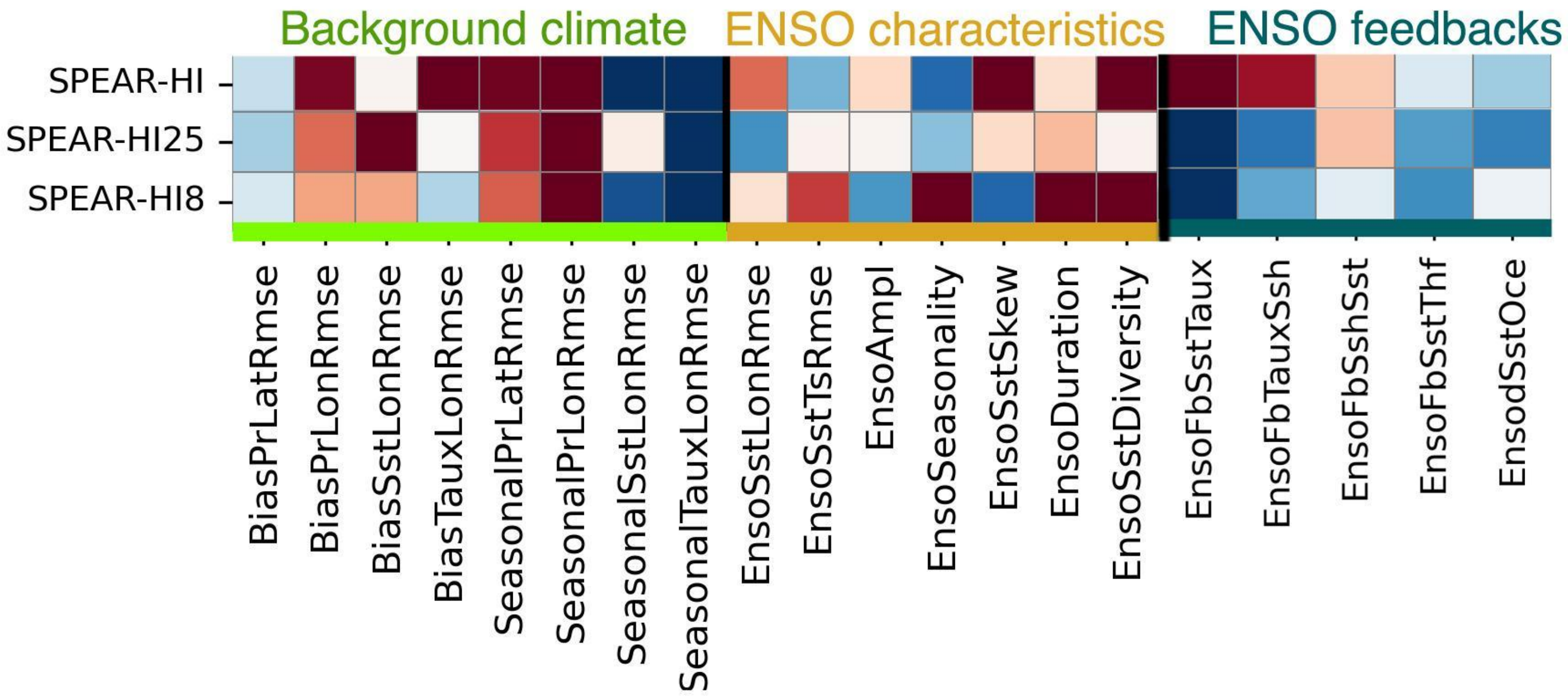
⇒ Improves simulation of tropical climate, and helps to attribute model biases.





Credit: Andrew Wittenberg (GFDL)

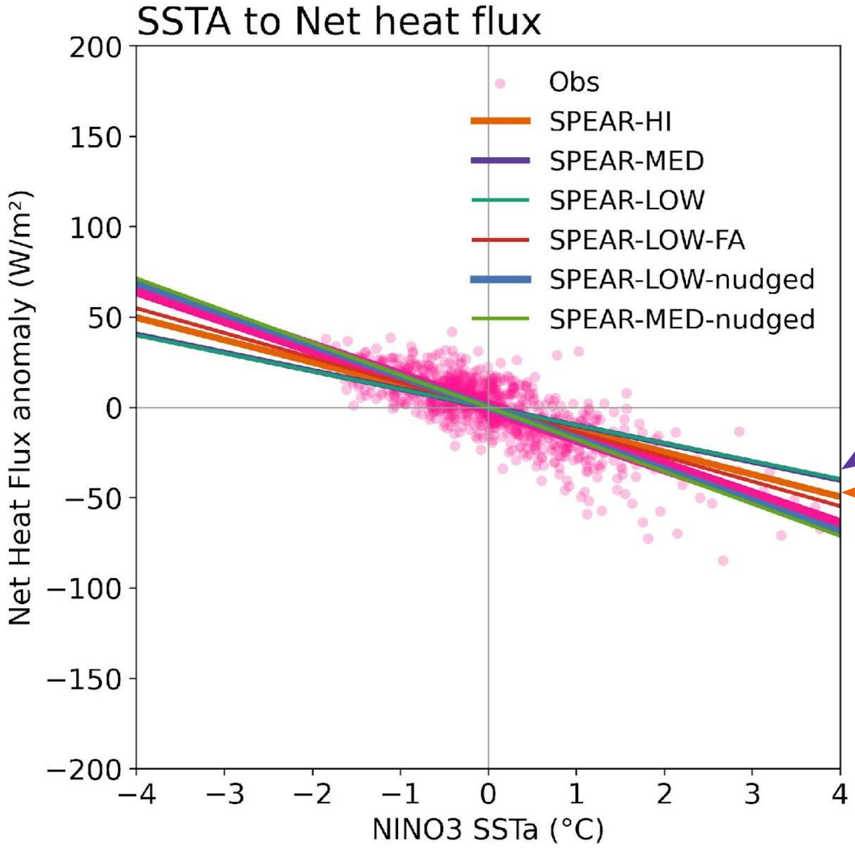
Ocean resolution: SPEAR-HI, -HI25, -HI8



Note that it is not an apples-to-apples comparison because:

SPEAR-HI25 (30members; Historical+SSP245): 2010-2084
SPEAR-HI8 (1990Control): Years 20-94

EN in reality
Warm SST →
↑ clouds → ↓ SW
↑ evaporation → ↑ latent heat loss
Net effect: ocean **loses heat**
→ **Negative feedback (thermal damping)**



EN in the SPEAR-HI
(under-damp system; i.e., too weak net heat flux feedback

With overestimated rainfall (model mean state bias)

- Convection is **too strong**
- Clouds respond **too strongly** to SST
- Evaporation responds **too strongly** to SST
- ⇒ **|∂Q′/∂T′| increases** (good if the model was under-damped)

Excessive heat loss for a given warm SST anomaly
El Niño is overdamped (too weak in -HI)

SPEAR-LOW and -MED

SPEAR-HI (closer to obs)

Mean state (Climatology)

1.double_ITCZ_bias: Meridional structure of time-mean PR in the eastern Pacific

Computes the meridional root mean square error (RMSE) of eastern Pacific (15°S-15°N) climatological (time and zonal 150-90°W average) precipitation (PR) between model and observations

2. eq_PR_bias: Zonal structure of time-mean PR in the equatorial Pacific

Computes the zonal root mean square error (RMSE) of equatorial Pacific (150°E-90°W) climatological (time and meridional 5°S-5°N average) precipitation (PR) between model and observations

3. eq_SST_bias: Zonal structure of time-mean SST in the equatorial Pacific

Computes the zonal root mean square error (RMSE) of equatorial Pacific (150°E-90°W) climatological (time and meridional 5°S-5°N average) sea surface temperature (SST) between model and observations

4. eq_Taux_bias: Zonal structure of time-mean Taux in the equatorial Pacific

Computes the zonal root mean square error (RMSE) of equatorial Pacific (150°E-90°W) climatological (time and meridional 5°S-5°N average) zonal wind stress (Taux) between model and observations

Mean state (Seasonal cycle)

5. double_ITCZ_sea_cycle: Meridional structure of the std of the seasonal cycle of precipitation in the eastern Pacific

Computes the meridional root mean square error (RMSE) of eastern Pacific (15°S-15°N) amplitude (standard deviation) of the mean annual cycle (zonal 150-90°W average) precipitation (PR) between model and observations (the time averaging creates a 12-month climatological time seres from which the standard deviation is computed)

6. eq_PR_sea_cycle: zonal structure of the std of the seasonal cycle of PR in the equatorial Pacific
Computes the zonal root mean square error (RMSE) of of equatorial Pacific (150°E-90°W) amplitude (standard deviation) of the mean annual cycle (zonal 5°S-5°N average) precipitation (PR) between model and observations
7. eq_SST_sea_cycle: Zonal structure of the std of the seasonal cycle of SST in the equatorial Pacific

Computes the zonal root mean square error (RMSE) of of equatorial Pacific (150°E-90°W) amplitude (standard deviation) of the mean annual cycle (zonal 5°S-5°N average) sea surface temperature (SST) between model and observations

8. eq_Taux_sea_cycle: Zonal structure of the std of the seasonal cycle of Taux in the equatorial Pacific

Computes the zonal root mean square error (RMSE) of of equatorial Pacific (150°E-90°W) amplitude (standard deviation) of the mean annual cycle (zonal 5°S-5°N average) zonal wind stress (Taux) between model and observations

ENSO Characteristics

9. ENSO_pattern: Zonal structure of boreal winter SST anomalies in the equatorial Pacific

Computes the zonal root mean square error (RMSE) of equatorial Pacific (150°E-90°W) sea surface temperature anomalies (SSTA; meridional 5°S-5°N average) during boreal winter (December value smoothed with a 5-month triangular-weighted moving average) between model and observations

10. ENSO_lifecycle: Temporal evolution of SST anomalies in the central equatorial Pacific

Computes the temporal root mean square error (RMSE) of central equatorial Pacific sea surface temperature anomalies (SSTA; Niño3.4 average) for 6 years centered on ENSO peak between model and observations

11. ENSO_amplitude: standard deviation of SST anomalies in the central equatorial Pacific

Computes the standard deviation of NINO3.4 sea surface temperature anomalies

12. ENSO_seasonality: ratio of boreal winter over spring's standard deviation of NINO3.4 SSTa

Computes the ratio of winter (NDJ, maximum variability in the observations) over spring (MAM, minimum variability in the observations) standard deviation of central equatorial Pacific sea surface temperature anomalies (SSTA; Niño3.4 average)

ENSO Characteristics

13. ENSO_asymmetry: Skewness of NINO3.4 SST anomalies

Computes the skewness of NINO3.4 sea surface temperature anomalies (SSTA)

14. ENSO_duration: Duration of SST anomalies in the central equatorial Pacific

Computes the number of months during which NINO3.4 SSTA lasts. It is based on the central equatorial Pacific SSTA (Niño3.4 average) during 6 years centered on ENSO peak between model and observations

15. ENSO_diversity: Diversity of zonal location of the maximum SSTA in the equatorial Pacific

Computes the interquartile range (IQR) of the distribution of zonal location of the maximum (minimum) sea surface temperature anomalies (SSTA) in the equatorial Pacific (meridional 5°S-5°N average) during boreal winter (December value smoothed with a 5-month triangular-weighted moving average) during El Niño (La Niña) events

ENSO feedback processes

16. SST-Taux_feedback: Coupling between SSTA in the eastern equatorial Pacific and TauxA in the western equatorial Pacific
Computes zonal wind stress anomalies (TauxA) in the western equatorial Pacific (Niño4 average) regressed onto SSTA in the eastern equatorial Pacific (Niño3 average).

17. Taux-SSH_feedback: Coupling between TauxA in the western equatorial Pacific and SSHA in the eastern equatorial Pacific
Computes sea surface height anomalies (SSHA) in the eastern equatorial Pacific (Niño3 average) regressed onto TauxA in the western equatorial Pacific (Niño4 average).

18. SSH-SST_feedback: Coupling between SSHA and SSTA in the eastern equatorial Pacific
Computes SSHA regressed onto SSTA in the eastern equatorial Pacific (Niño3 average).

19. SST-NHF_feedback: Coupling between SSTA and NHFA in the eastern equatorial Pacific
Computes net surface heat flux anomalies (NHFA; sum of latent and sensible heat fluxes and longwave and shortwave radiations) regressed onto SSTA in the eastern equatorial Pacific (Niño3 average).

20. ocean_driven_SST: SSTA caused by anomalous ocean circulation in the eastern eqPacific
Computes the amount of cooling (warming) by anomalous ocean circulation needed to generate surface temperature anomalies (SSTA) of -1 (1) °C during La Nina (El Nino) events in the eastern equatorial Pacific (Niño3 average).
Formula: dSST_oce = dSST – dSST_nhf