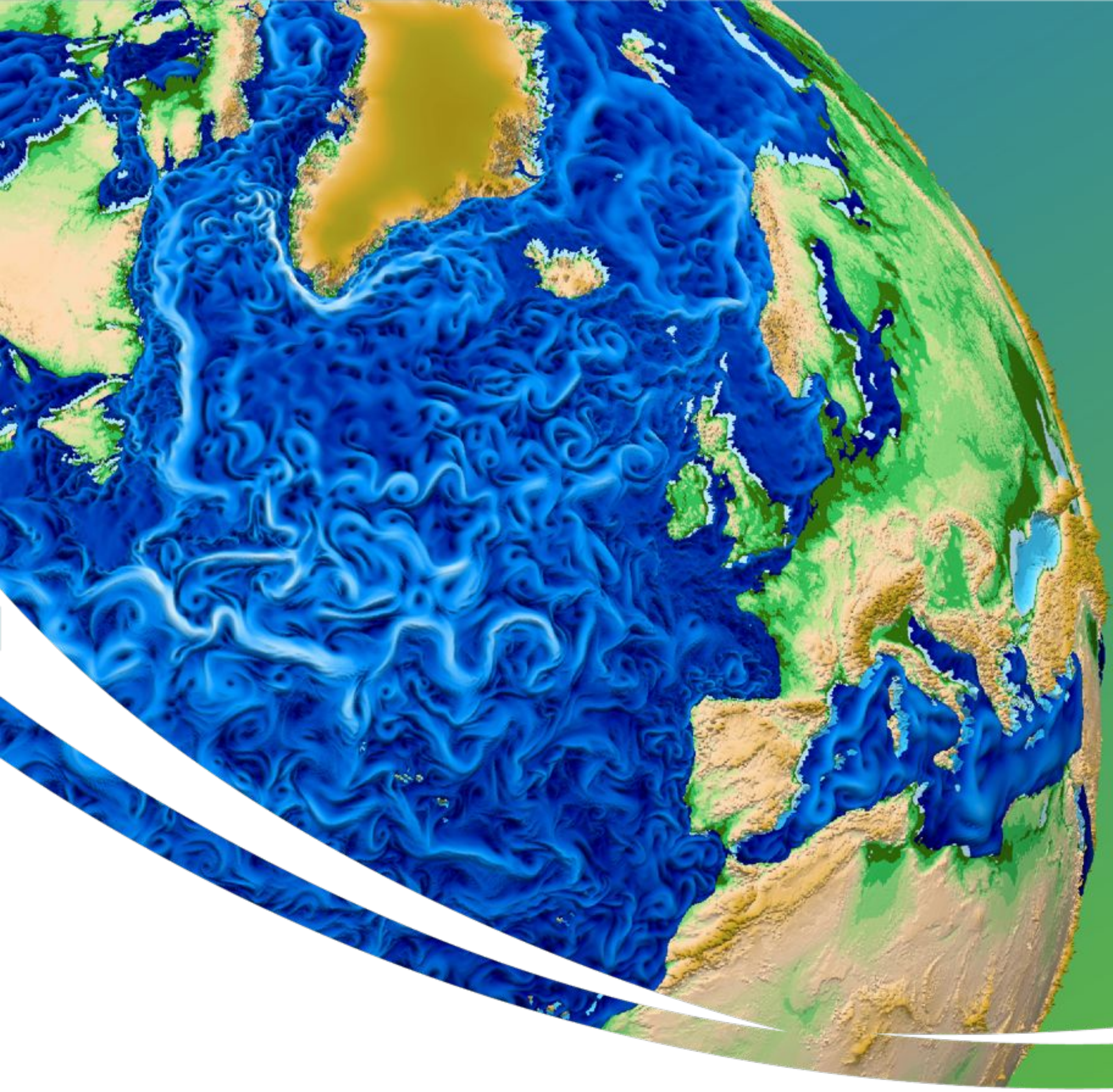


Multi-Pronged Approaches for Addressing Model Biases Relevant to S2S Predictions

L. Ruby Leung

Pacific Northwest National Laboratory

U.S. DEPARTMENT OF ENERGY

E3SM Energy Exascale Earth System Model

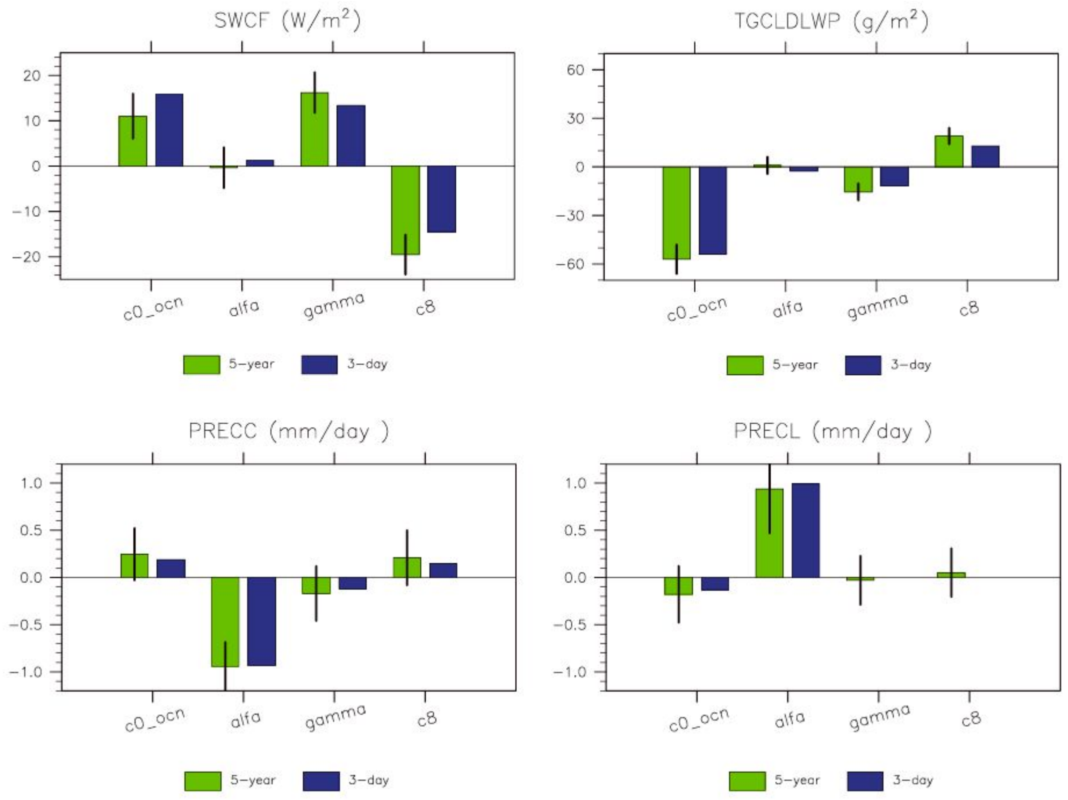
Addressing three sources of model biases

- The climate mean states
- The S2S modes of variability (MoV)
- Relationship between MoV and surface climate and extreme events

Through model calibration and diagnostics

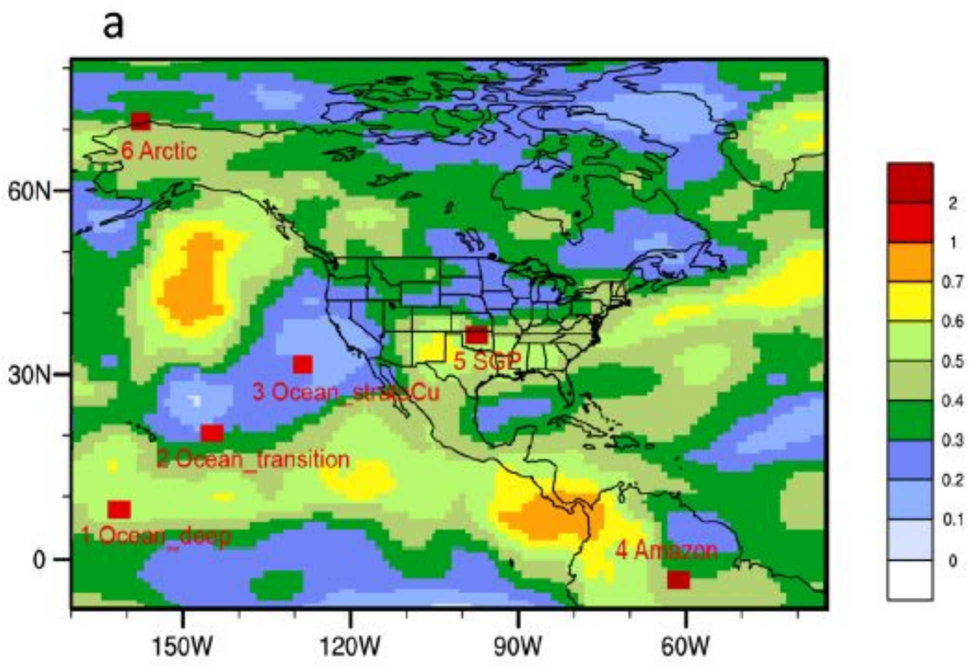
Model calibration strategy: short perturbed physics experiments (PPE)

Target fast response of physical processes (e.g., turbulence, microphysics, shallow and deep convection) – select 18 tunable parameters

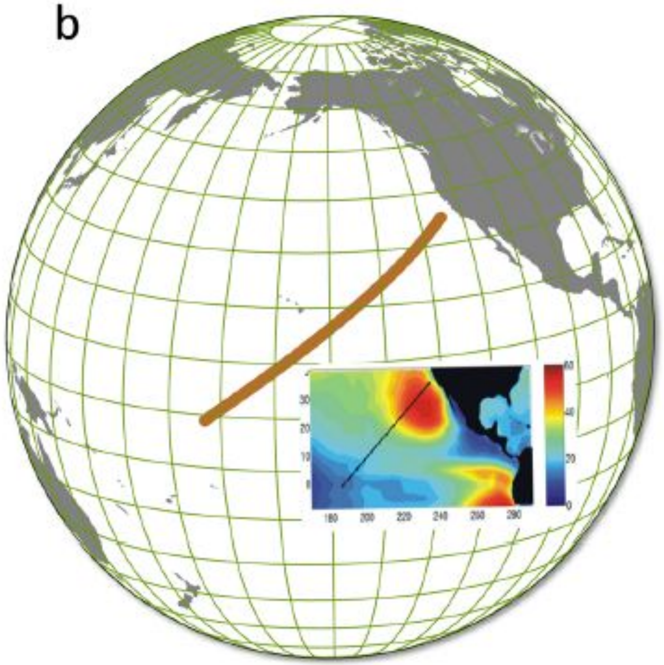


(Qian et al. 2019 JGRA)

Model calibration strategy: diverse climate and cloud regimes



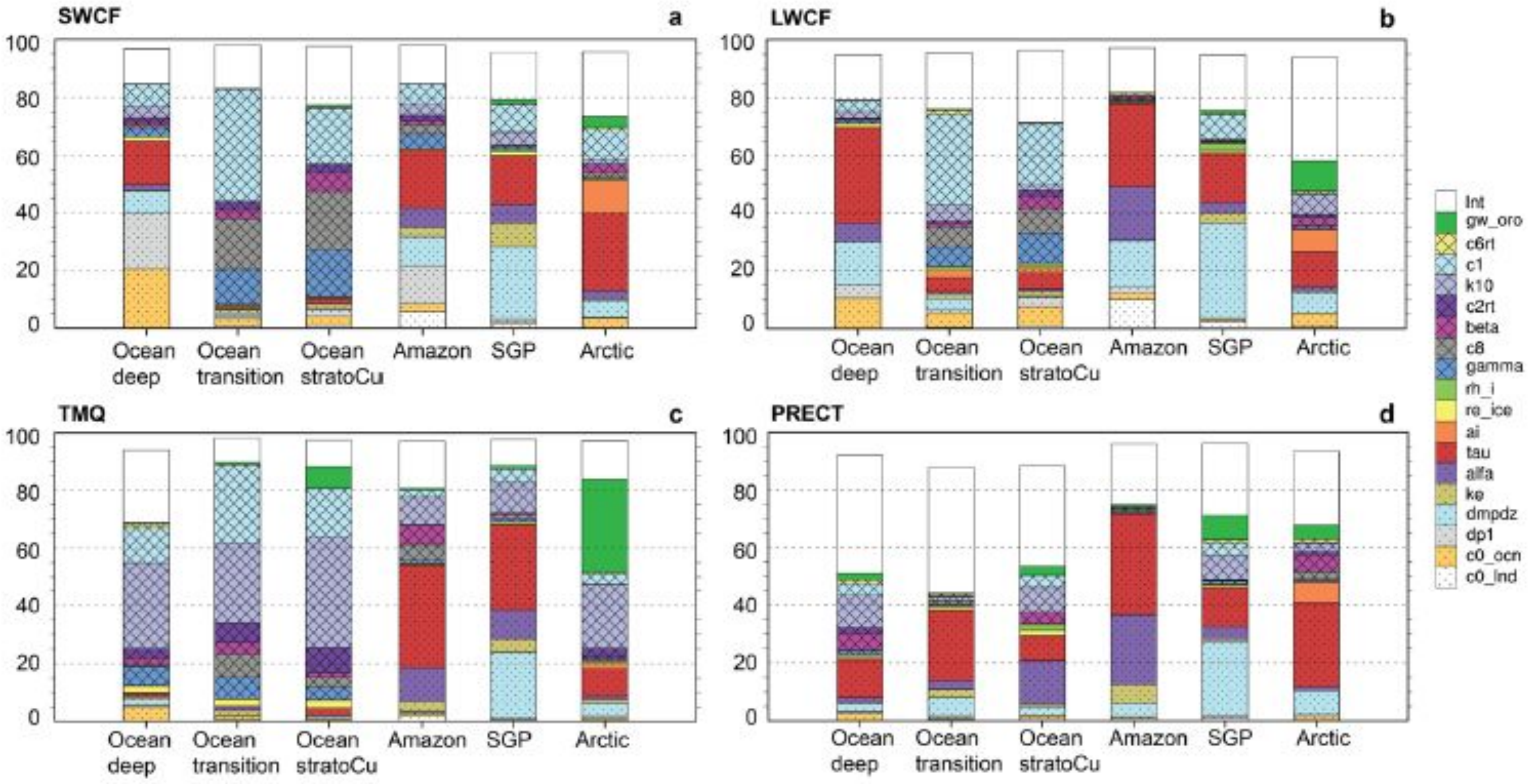
Six regions over land and ocean



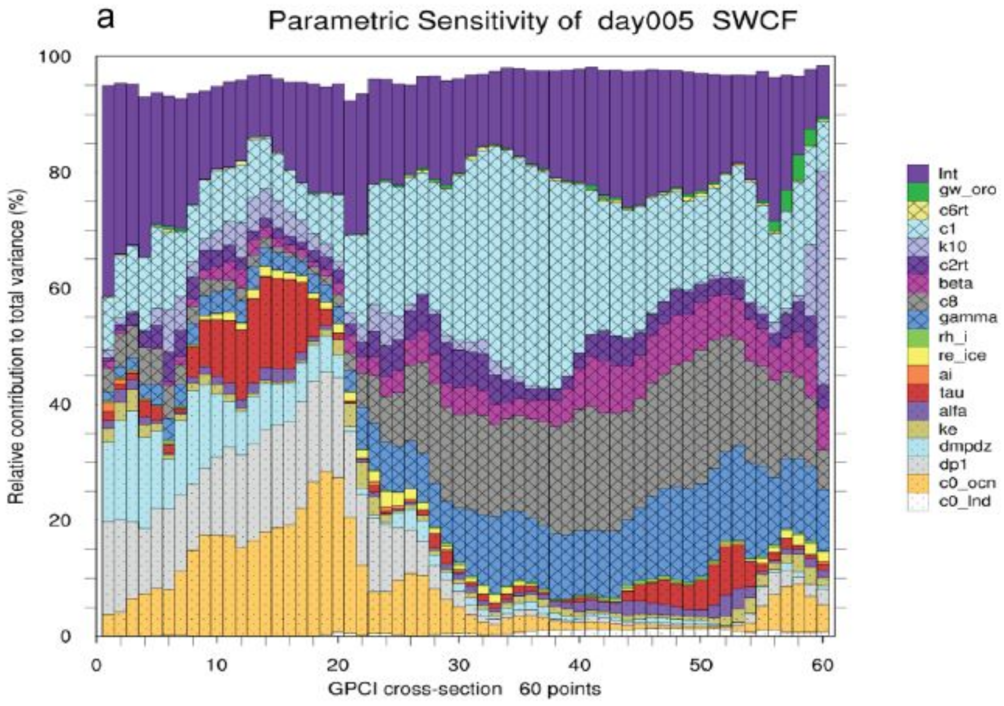
GEWEX Cloud System Study (GCSS) Pacific Cross-section Intercomparison (GPCI)

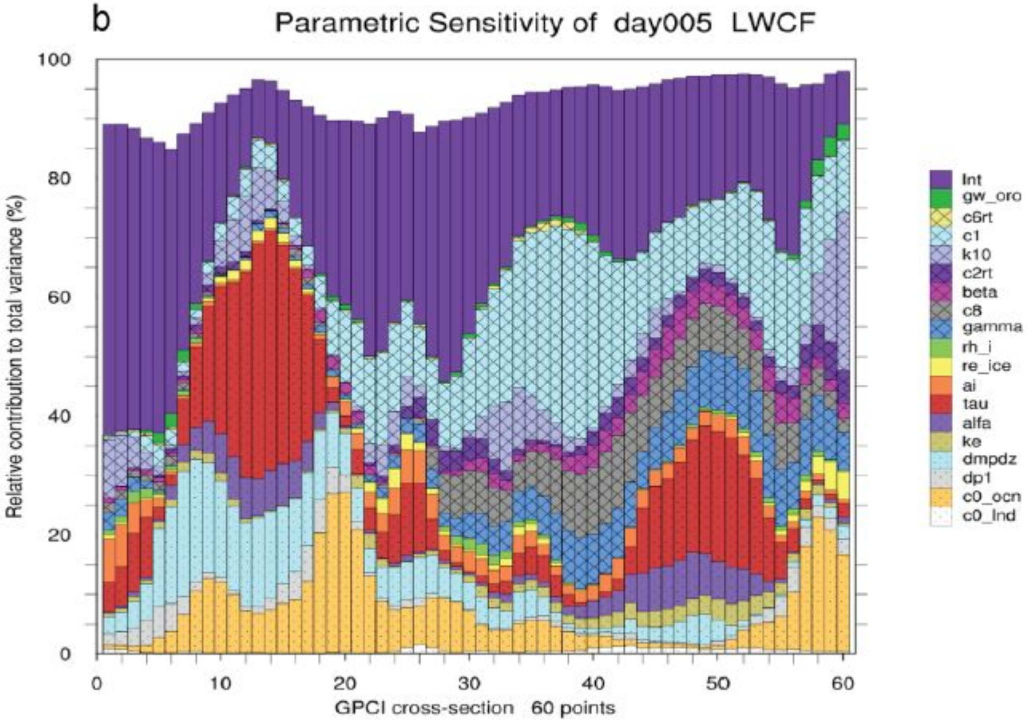
(Qian et al. 2024 Clim. Dyn.)

Relative contributions of different parameters to the total variance in the six regions on Day 5



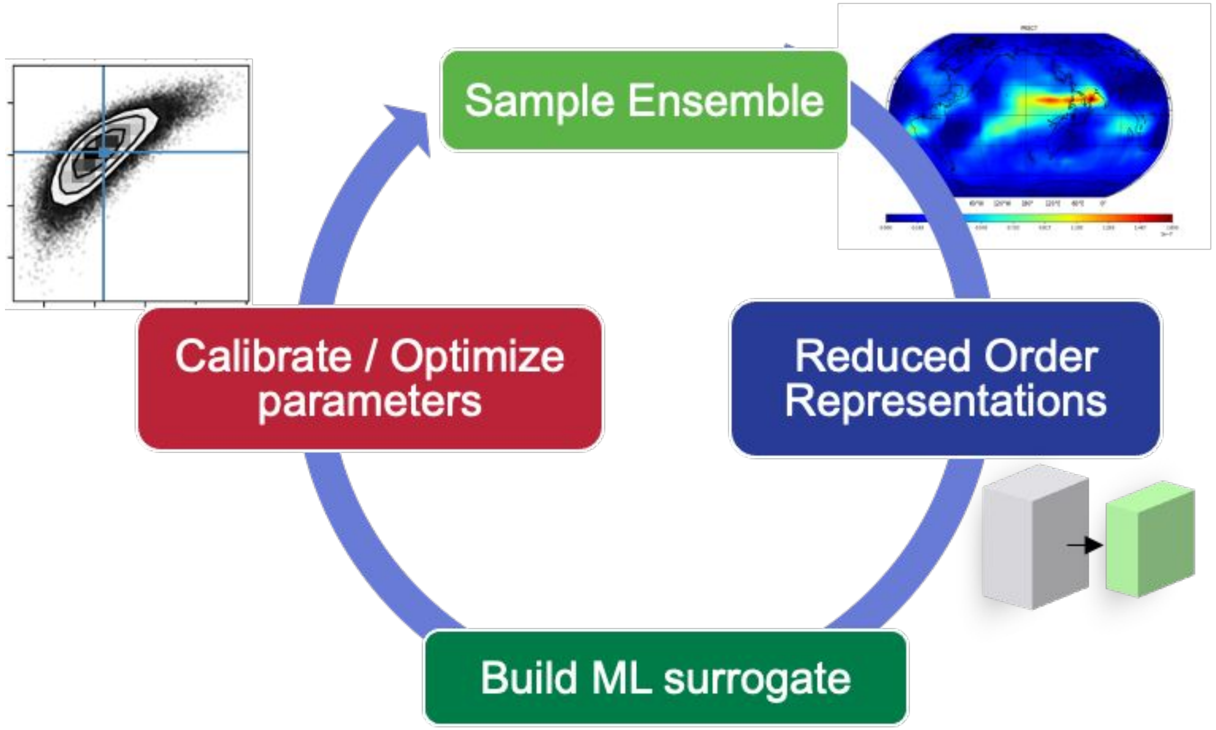
Relative contributions of different parameters to the total variance along the GPCI transect on Day 5

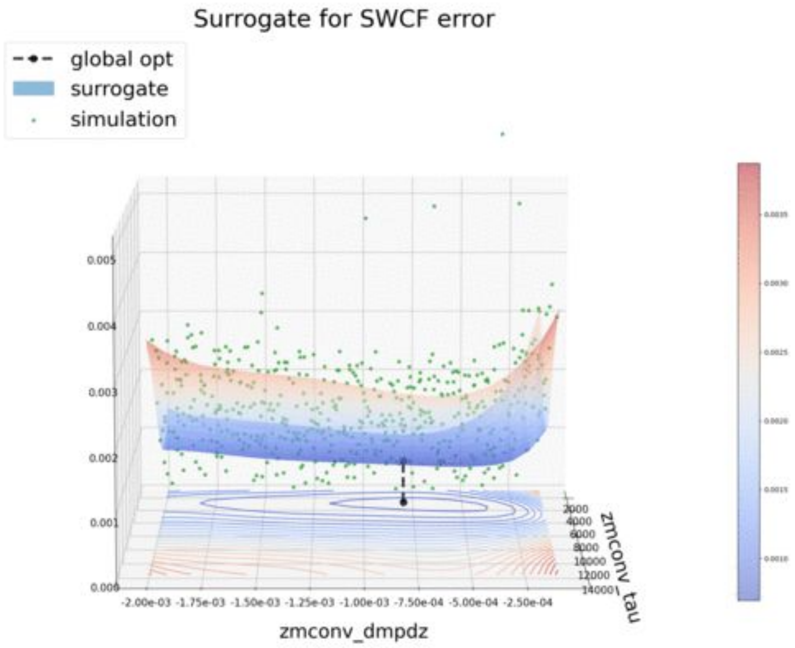




Using ML Surrogates as part of an efficient calibration process

Lays the groundwork for rigorous Bayesian UQ





POC: Wagman

Using ML Surrogates as part of an efficient calibration process

Results for E3SMv2:

- 5 parameters
- 350 ensemble members
- 44 climate metrics

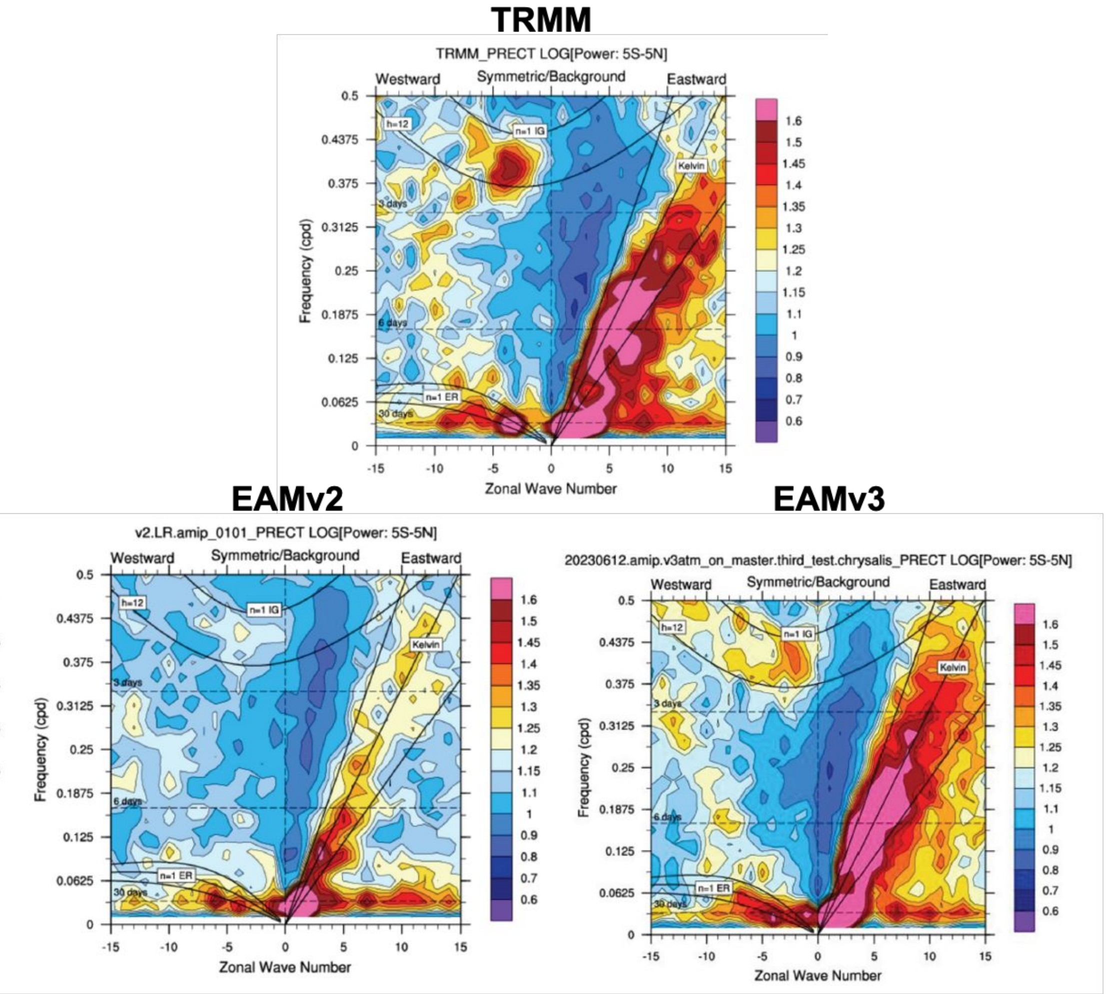
- Most metrics improved (green boxes in table)
- Offline optimization calculations are quick

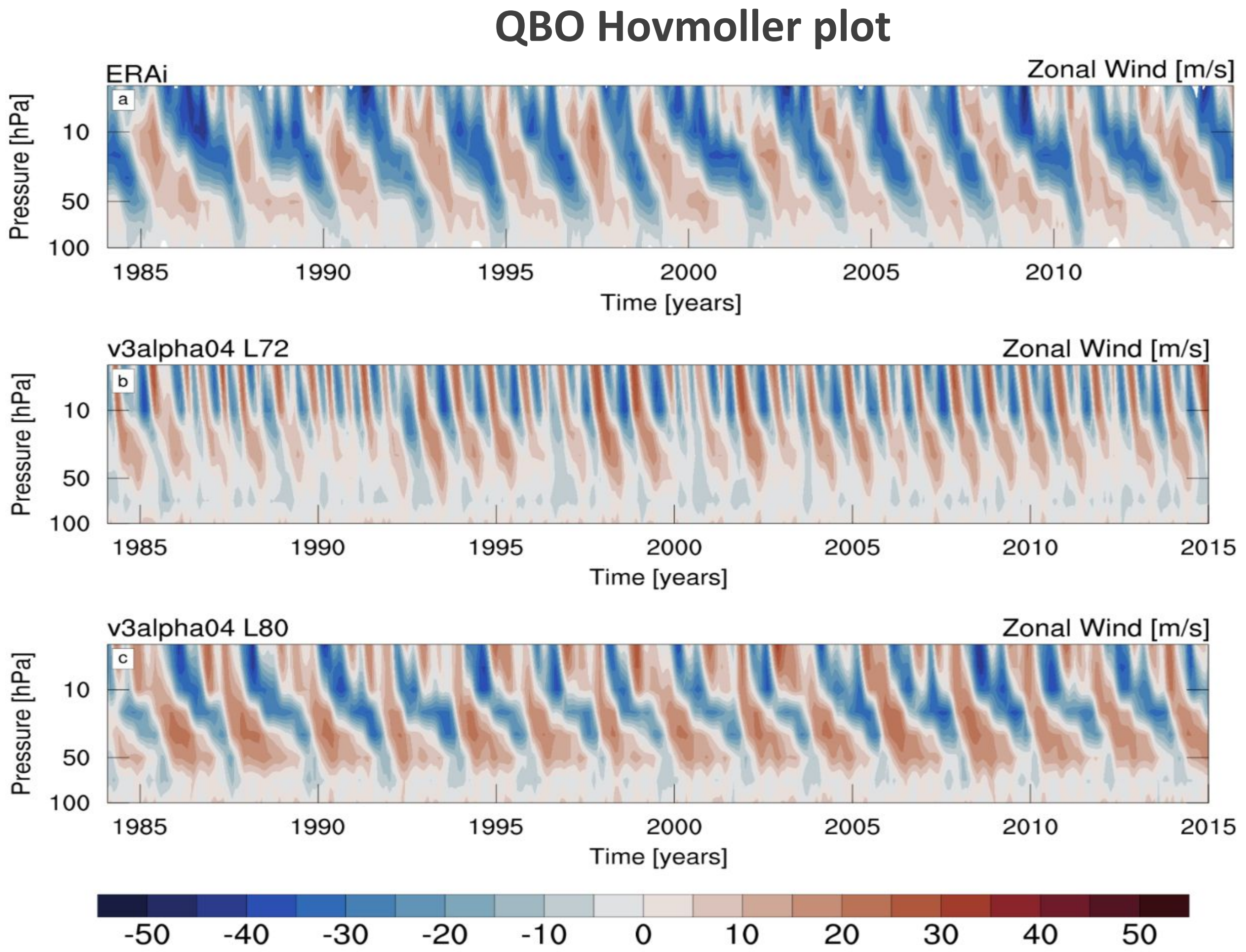
Percent change of 44 bias metrics at calibrated parameter values compared to previous tuning

Variable	DJF	MAM	JJA	SON	Avg.
LWCF	9.7	−1.3	0.4	10.0	4.7
PRECT	9.5	4.1	−0.3	11.8	6.3
PSL	4.3	−6.9	−5.3	−18.0	−8.6
RELHUM	−1.7	0.3	1.9	0.4	0.2
SWCF	5.1	−0.3	−6.2	2.0	0.1
T	−0.3	−3.3	1.9	−4.0	−1.4
TREFHT	−7.2	−10.0	−2.5	−10.3	−7.5
U	1.4	−10.6	−6.7	−10.8	−6.7
U200	7.4	−12.8	−18.0	−7.3	−4.0
U850	5.7	−11.8	−16.1	0.7	−5.4
Z500	4.0	−9.8	−7.1	−15.0	−2.7
Average	2.7	−5.7	−5.3	−2.4	−2.7

POC: Wagman

Improved tropical variability in v3

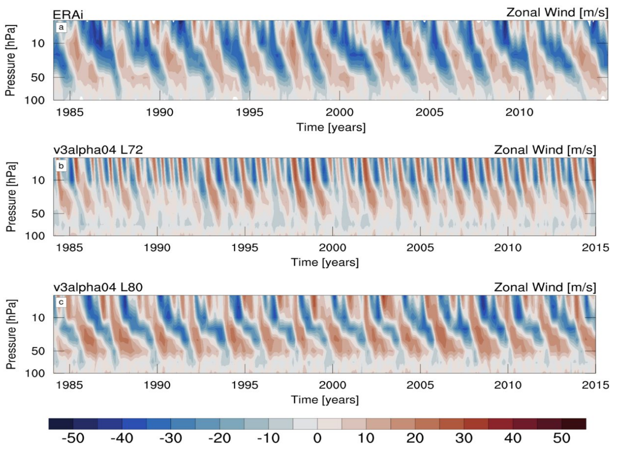


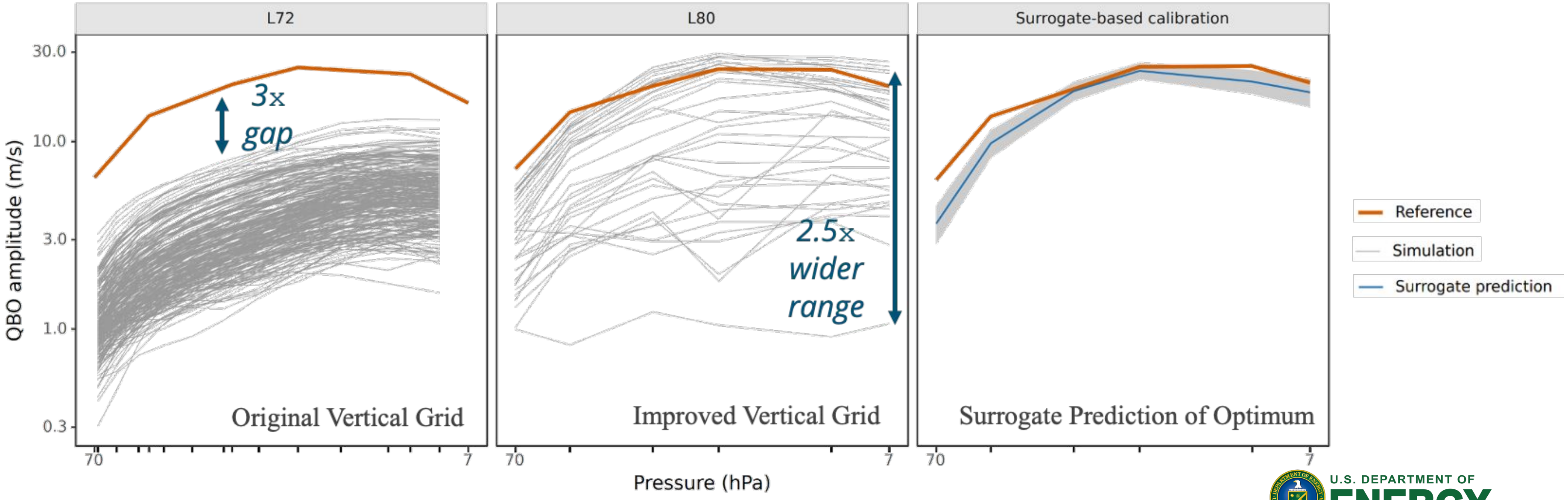


E3SM
Energy Exascale
Earth System Model

Improving QBO in E3SMv3

- The QBO bias was not appreciably improved by any member of the perturbed parameter ensemble, pointing to the need for model enhancements.
- Redesigning the vertical layers made E3SM more responsive to parameter tuning. The new grid was added in time for its release as part of E3SMv3.
- Our workflow revealed an opportunity to expose new physics parameters and further refine the vertical grid for a better QBO.



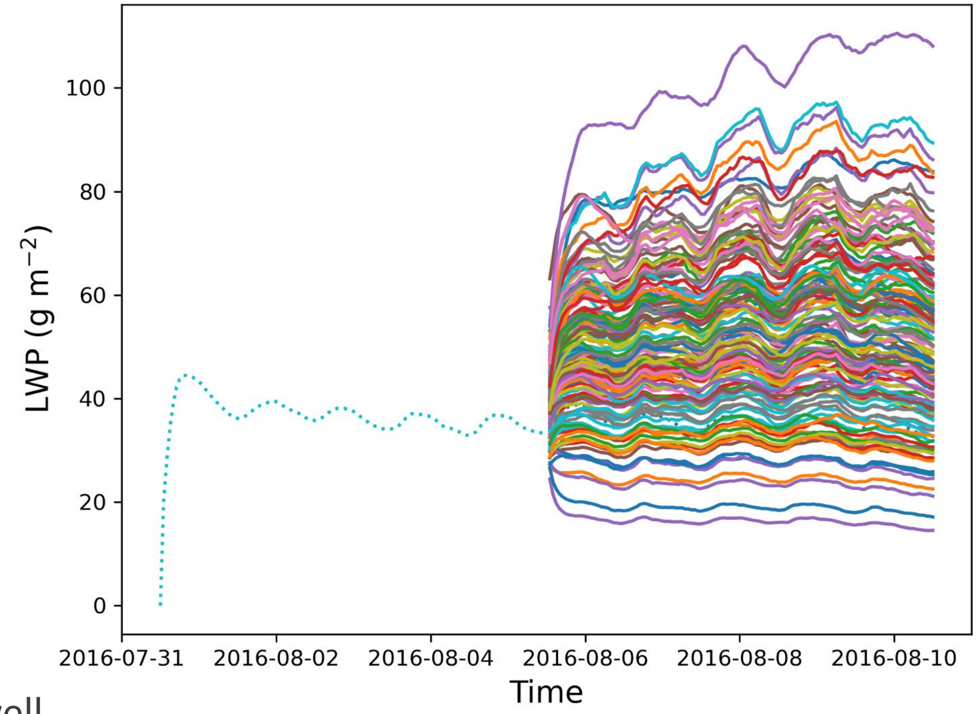


POC: Benedict

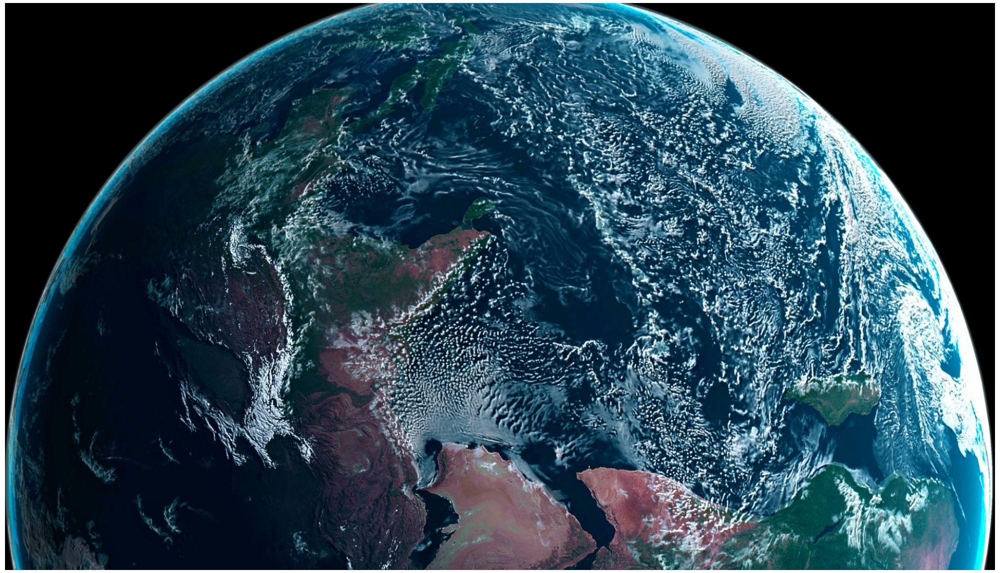
U.S. DEPARTMENT OF ENERGY

Automated calibration of SCREAM at 3 km

- SCREAM needs a full exascale machine to run at 1 SYPD; how can we afford to calibrate?
- **Short 2-5 day forecasts** can capture emergent cloud properties, and are affordable
- We calibrated 16 atm physics parameters to minimize a cost function of 30 cloud metrics using **300 SCREAM runs** on Frontier.



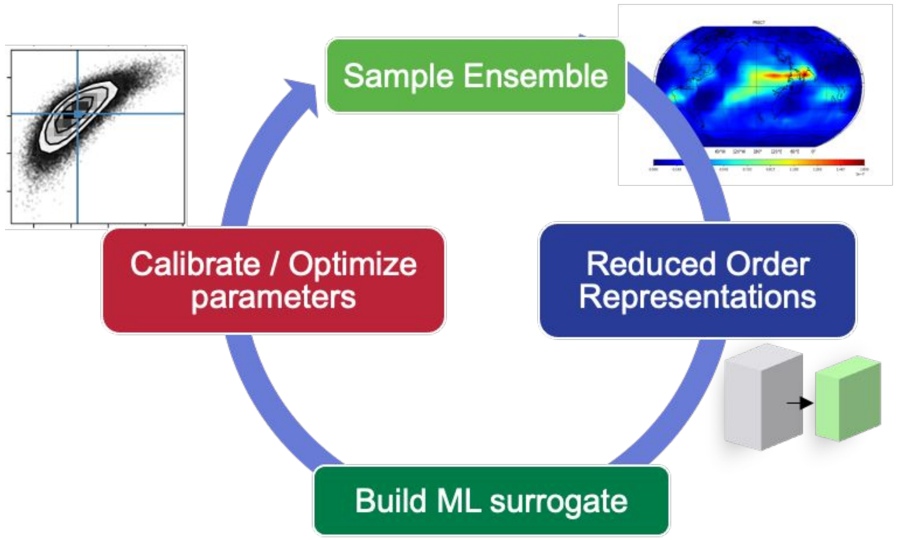
Rendering of clouds from a SCREAM simulation



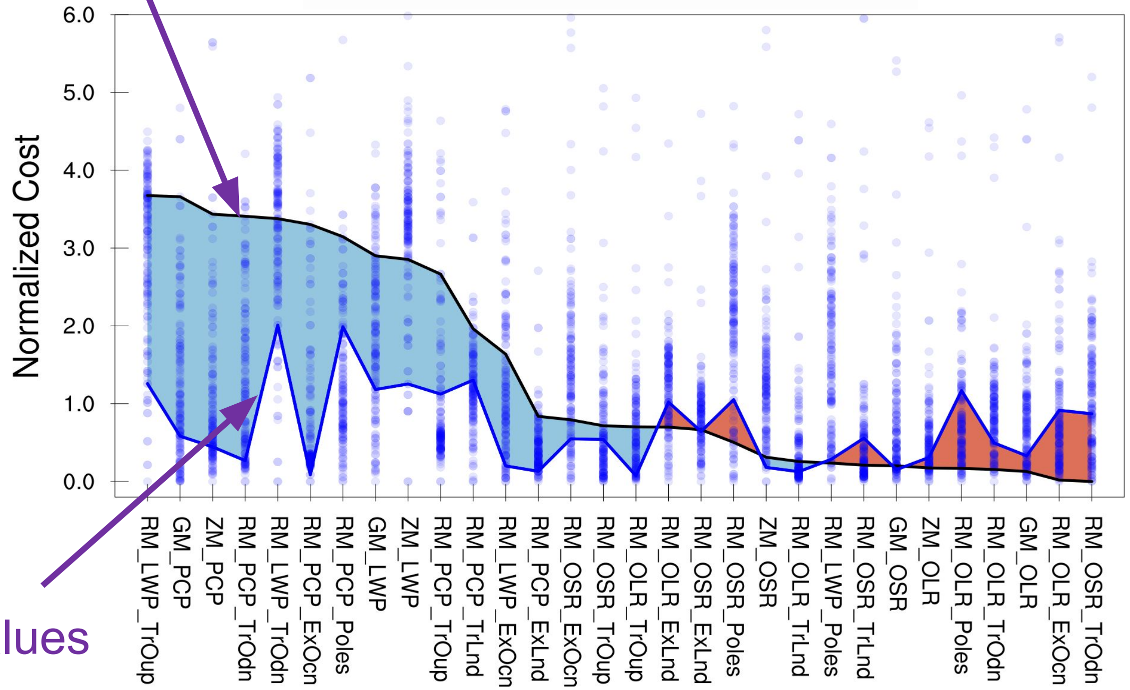
Roeber and Kuhn (NVIDIA)

POC: Caldwell

Automated calibration of SCREAM at 3 km



Default parameter values



Optimized parameter values

US Climate Modeling Summit project: Is Better Representation of Modes of Variability Related to Reduced Biases and Better Simulations of Extreme Events in US Climate Models?

PI: Gokhan Danabasoglu

Methods

The Climate Variability Diagnostics Package for large ensembles is used to evaluate climate modes in observations and models.

The package leverages the opportunities provided by LEs to:
1) remove the forced response to avoid aliasing of the forced response
2) quantify the ensemble spread and average in mode benchmarks
3) confidently assess inter-model contrasts and inter-generational changes in model fidelity,
4) estimate the limits of the observational record in sampling climate modes (e.g. finding that intrinsic noise in the IPO limits the utility of 100-yr records in model evaluation).

Ensemble Sizes:
CESM1 (40), CESM2 (100), E3SM1 (14), E3SM2 (21), GISS E2.1-G (46)
GISS E2.1-H (25), GISS E2.2-G (10), GISS E2.2-H (5), GFDL CM4 (20)
GFDL ESM2M (30), GFDL SPEAR (30)

(Fasullo et al. 2024 JCLIM)

NCAR | COMMUNITY EARTH SYSTEM MODEL CESM

Home About What We Do Models Working Groups Events News

Climate Variability Diagnostics Package for Large Ensembles (CVDP-LE)

Version Information Current Version: v1.0.0

The Climate Variability Diagnostics Package for Large Ensembles (CVDP-LE) developed by NCAR's Climate Analysis Section is an automated analysis tool and data repository for exploring internal and forced contributions to climate variability and change in coupled model "initial-condition" Large Ensembles and observations.

The package computes a wide range of modes of interannual-to-multidecadal variability in the atmosphere, ocean and cryosphere, as well as long-term trends and key indices of global and regional climate. Diagnostics include the ensemble-mean (i.e., forced response) and ensemble-spread (i.e., internal variability) of each model, as well as quantitative metrics comparing the models to observations. All diagnostics and metrics are saved to a data repository for later use and analysis.

The CVDP-LE User's Guide provides general background on initial-condition Large Ensembles, detailed documentation of all diagnostics and metrics in the package, and guidance on interpreting the results. Instructions for downloading and running the CVDP-LE are provided on the Code page and readme file, respectively.

The CVDP-LE can be applied to any suite of observational data, model simulations and time periods specified by the user. A few examples of CVDP-LE applications to the CESM2 Large Ensemble, the Multi-Model Large Ensemble Archive and the CMIP6 archive are linked below; additional comparisons including netCDF files of CVDP-LE calculations can be found in the Data Repository.

Key Points

The fidelity of US Climate Models in simulating major modes of variability and their teleconnections has generally improved across generations.

Recently produced large ensembles have been a key tool for quantifying inter-model contrasts and estimating the intrinsic noise of climate modes.

Pattern correlations against observations of the major modes of variability by each US Center has generally improved across model generation.

Ensemble-Average Mean Scores

CESM1 (0.82) < CESM2 (0.87)

E3SM1 (0.84) < E3SM2 (0.86)

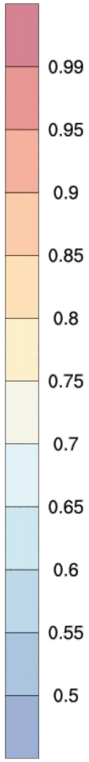
GISS E2.1 G (0.77) < GISS E2.2 G (0.80)

GFDL CM4 (0.82) ≅ GFDL ESM2M (0.82) < SPEAR (0.87)

(Fasullo et al. 2024 JCLIM)

Pattern Correlations

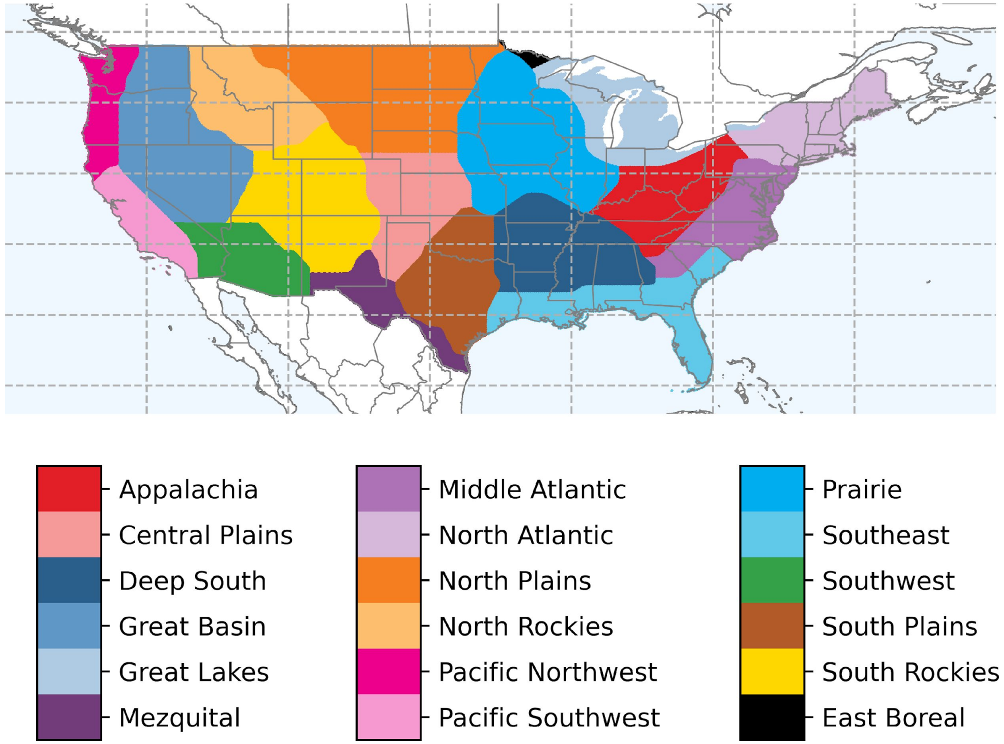
	ENSO TAS (DJF^{+1})	ENSO PSL (DJF^{+1})	El Niño Hovmöller	La Niña Hovmöller	AMV' Low-Pass	PDV'	NAO (JFM)	PNA (DJF)	SAM (DJF)	SST std dev (Ann)	PR std dev (Ann)	Mean Score
CESM1 (10%)	0.37	0.79	0.91	0.88	0.23	0.85	0.85	0.90	0.84	0.68	0.76	0.81
CESM1 (Avg)	0.45	0.82	0.93	0.91	0.39	0.88	0.88	0.92	0.86	0.70	0.76	0.82
CESM1 (90%)	0.52	0.85	0.94	0.93	0.52	0.89	0.90	0.93	0.88	0.71	0.77	0.83
CESM2 (10%)	0.64	0.81	0.93	0.90	0.37	0.88	0.90	0.94	0.86	0.76	0.83	0.86
CESM2 (Avg)	0.70	0.86	0.94	0.91	0.53	0.90	0.92	0.96	0.88	0.77	0.84	0.87
CESM2 (90%)	0.73	0.89	0.95	0.93	0.67	0.91	0.94	0.97	0.90	0.78	0.84	0.88
CESM2smbb (10%)	0.63	0.83	0.93	0.90	0.35	0.87	0.90	0.94	0.85	0.75	0.83	0.86
CESM2smbb (Avg)	0.69	0.86	0.94	0.91	0.52	0.90	0.92	0.95	0.88	0.77	0.84	0.87
CESM2smbb (90%)	0.73	0.90	0.95	0.92	0.65	0.91	0.94	0.96	0.90	0.78	0.84	0.88
E3SM1 (10%)	0.46	0.65	0.86	0.82	0.27	0.59	0.90	0.91	0.94	0.64	0.81	0.82
E3SM1 (Avg)	0.63	0.76	0.89	0.85	0.37	0.79	0.94	0.94	0.96	0.66	0.82	0.84
E3SM1 (90%)	0.72	0.84	0.91	0.88	0.53	0.86	0.97	0.96	0.96	0.67	0.82	0.85
E3SM2 (10%)	0.38	0.81	0.93	0.88	0.39	0.87	0.88	0.91	0.96	0.61	0.83	0.85
E3SM2 (Avg)	0.50	0.86	0.94	0.91	0.50	0.89	0.91	0.94	0.97	0.63	0.83	0.86
E3SM2 (90%)	0.57	0.89	0.95	0.92	0.60	0.90	0.93	0.95	0.97	0.64	0.84	0.87
GISSE21G (10%)	0.63	0.67	0.73	0.69	0.26	0.73	0.85	0.79	0.88	0.67	0.72	0.75
GISSE21G (Avg)	0.66	0.76	0.76	0.72	0.41	0.79	0.89	0.86	0.91	0.68	0.73	0.77
GISSE21G (90%)	0.69	0.82	0.79	0.76	0.55	0.83	0.92	0.92	0.93	0.69	0.75	0.79
GISSE21H (10%)	0.61	0.71	0.73	0.79	0.29	0.77	0.90	0.83	0.92	0.55	0.60	0.78
GISSE21H (Avg)	0.69	0.80	0.80	0.83	0.41	0.80	0.93	0.89	0.93	0.58	0.63	0.80
GISSE21H (90%)	0.76	0.86	0.86	0.87	0.56	0.84	0.95	0.93	0.95	0.61	0.65	0.82
GISSE22G (10%)	0.51	0.71	0.82	0.77	0.06	0.70	0.48	0.64	0.93	0.67	0.74	0.76
GISSE22G (Avg)	0.64	0.82	0.86	0.83	0.37	0.79	0.87	0.86	0.94	0.68	0.76	0.80
GISSE22G (90%)	0.71	0.86	0.89	0.85	0.55	0.85	0.93	0.93	0.95	0.68	0.78	0.83
GISSE22H (10%)	0.51	0.71	0.82	0.77	0.12	0.71	0.88	0.80	0.94	0.67	0.77	0.76
GISSE22H (Avg)	0.62	0.80	0.85	0.82	0.30	0.77	0.90	0.86	0.95	0.68	0.77	0.80
GISSE22H (90%)	0.71	0.85	0.90	0.85	0.49	0.81	0.91	0.93	0.95	0.69	0.78	0.83
GFDLCM4 (10%)	0.33	0.83	0.86	0.78	0.49	0.81	0.90	0.91	0.93	0.39	0.77	0.81
GFDLCM4 (Avg)	0.41	0.87	0.88	0.80	0.61	0.84	0.93	0.93	0.94	0.43	0.77	0.82
GFDLCM4 (90%)	0.48	0.89	0.89	0.82	0.66	0.85	0.95	0.94	0.94	0.46	0.78	0.83
GFDLESM2M (10%)	0.55	0.73	0.83	0.70	0.33	0.79	0.91	0.89	0.94	0.62	0.75	0.81
GFDLESM2M (Avg)	0.60	0.79	0.85	0.77	0.46	0.82	0.94	0.93	0.95	0.66	0.76	0.82
GFDLESM2M (90%)	0.64	0.84	0.87	0.84	0.56	0.84	0.95	0.95	0.95	0.68	0.76	0.83
GFDLSPEAR (10%)	0.46	0.81	0.92	0.92	0.43	0.77	0.91	0.82	0.95	0.75	0.86	0.85
GFDLSPEAR (Avg)	0.55	0.86	0.94	0.94	0.54	0.83	0.94	0.89	0.96	0.77	0.87	0.87
GFDLSPEAR (90%)	0.62	0.89	0.95	0.95	0.62	0.86	0.96	0.93	0.96	0.78	0.87	0.88



Connecting model skill of MoV with ETCCDI (extreme indices)

- Goal: to use MoV time series of predict ETCCDI time series in each Bukovsky region
- 17 Bukovsky regions
- MoV and ETCCDI time series are detrended
- **Random Forest (RF) regression model**

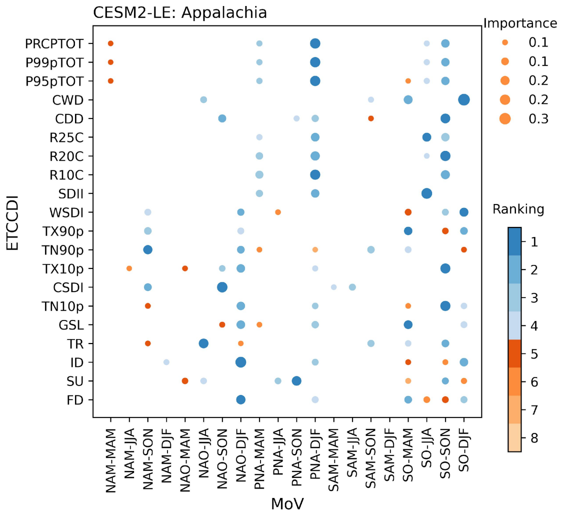
Bukovsky regions over CONUS

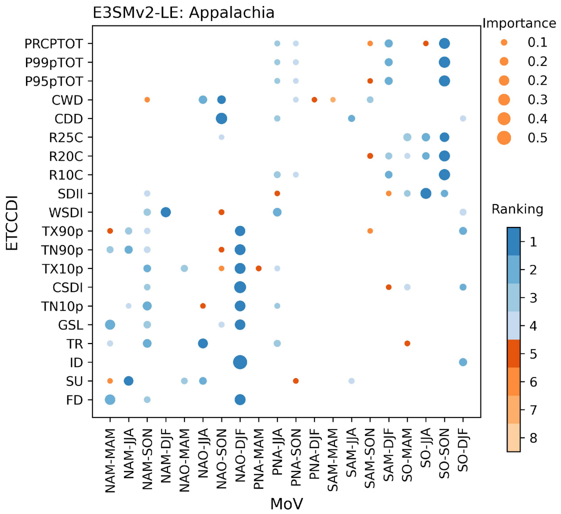


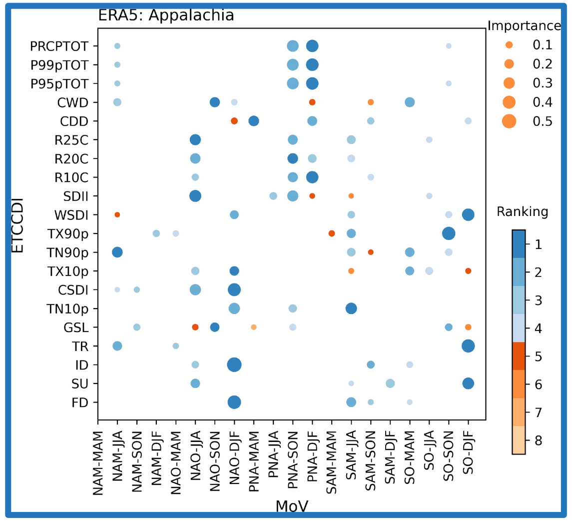
MoV	Description	Time series	Pattern
NAM	Northern Annular Mode	✓	✓
NAO	North Atlantic Oscillation	✓	✓
PNA	Pacific North American teleconnection pattern	✓	✓
SAM	Southern Annular Mode	✓	✓
SO	Southern Oscillation	✓	✓
NPO	North Pacific Oscillation	x	✓
PSA1	Pacific South American teleconnection pattern 1	x	✓
PSA2	Pacific South American teleconnection pattern 2	x	✓

Relative importance of different MoVs in predicting ETCCDI

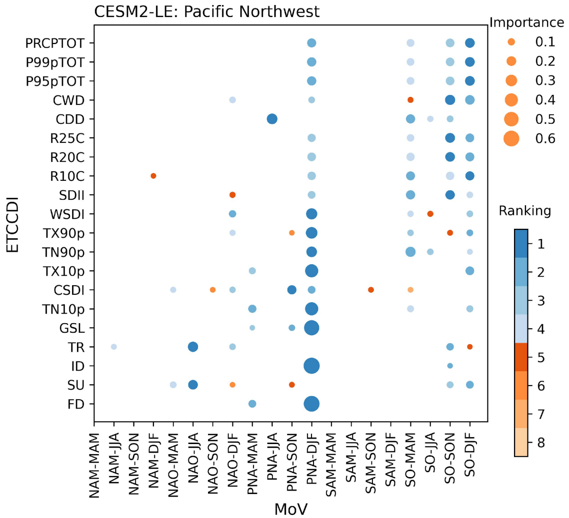
Appalachia

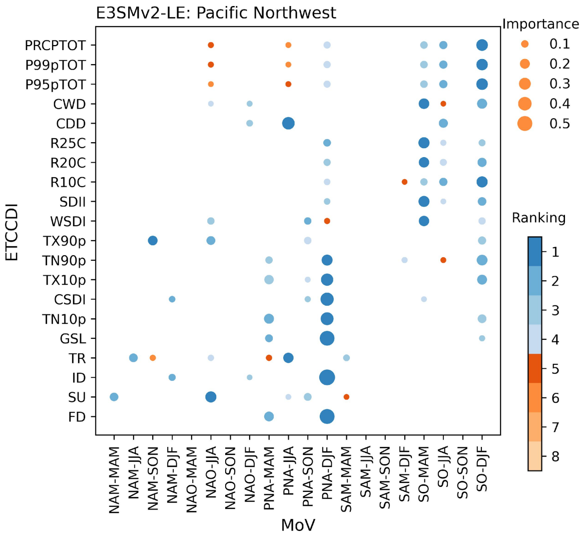


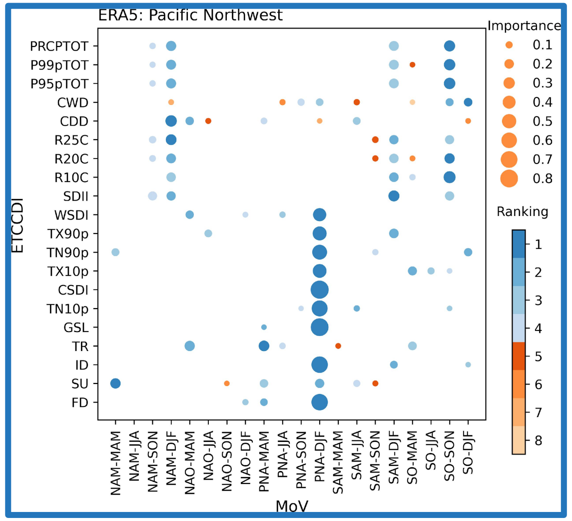




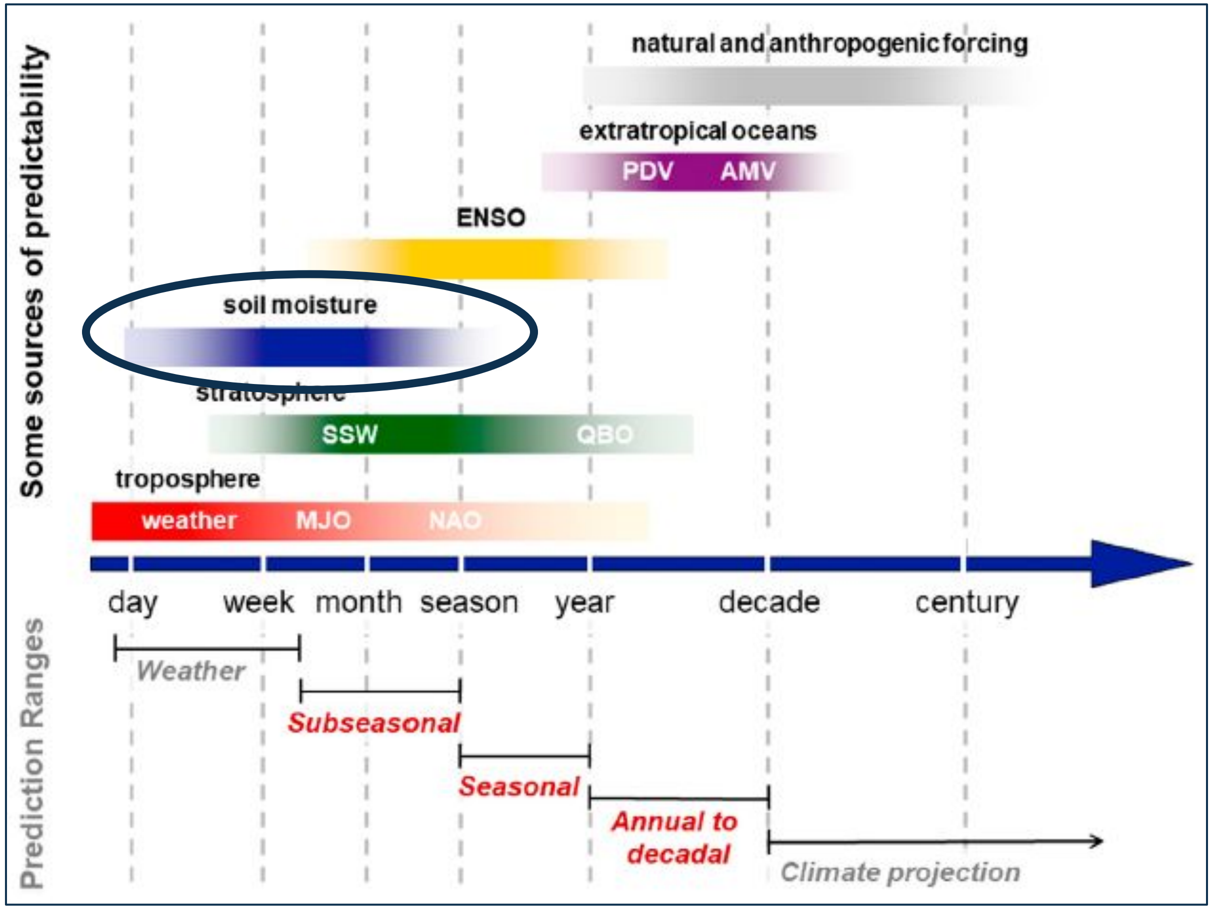
Pacific Northwest







Sources of predictability at different timescales



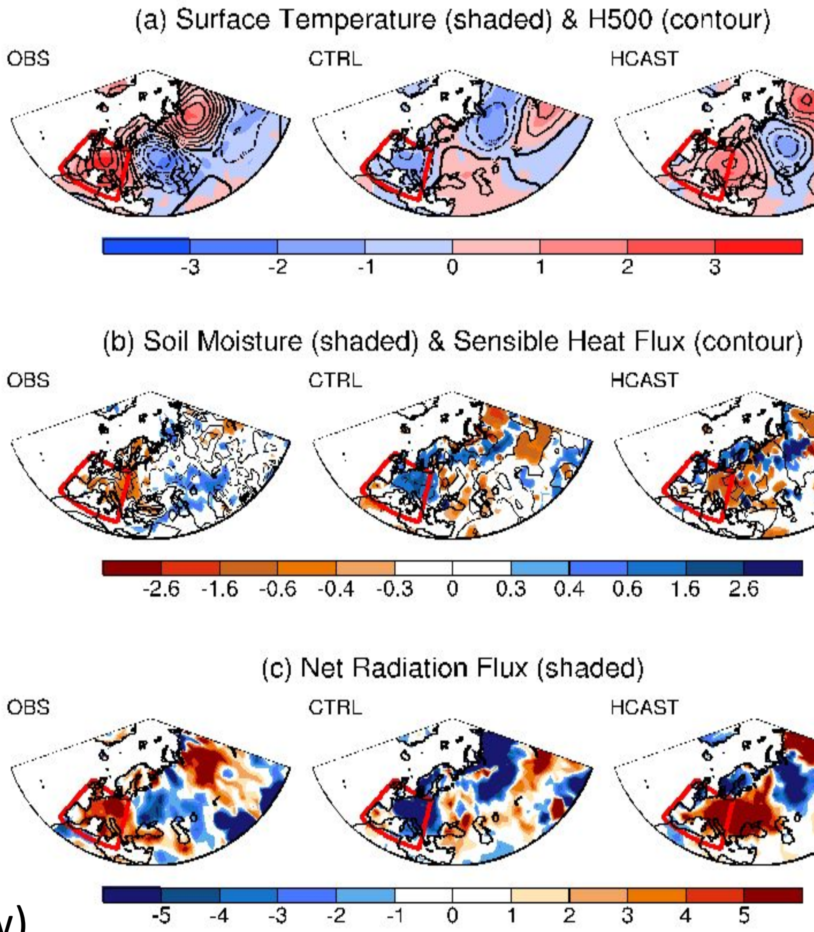
Could soil moisture and temperature provide predictability at interannual-to-decadal timescale through **longer memory land processes** and/or their **influence on ocean with longer memory**?

(Merryfield et al. 2020 BAMS)

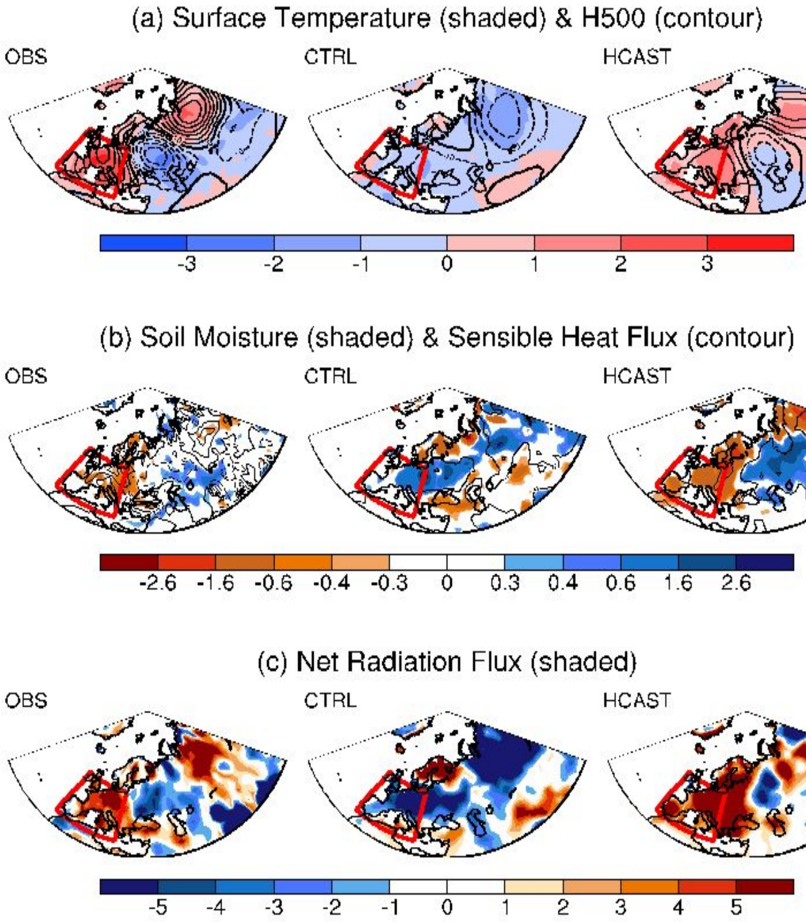
E3SM Energy Exascale Earth System Model

Hindcasts initialized in 2001 from a simulation where soil moisture/temperature are assimilated captures the 2003 European summer heatwave

FGOALS



E3SMv2



(Shi et al. in review)

Summary

- A robust model calibration strategy including use of AI/ML can reduce model biases in the mean states and MoV, though addressing model structural issues is critically important
- A short simulation strategy for model calibration makes it feasible to calibrate model parameters for global convection permitting models
- Leveraging and developing model diagnostics and metrics is important to characterize model biases as targets for improvements
- Soil moisture/temperature could provide an important source of predictability at S2S timescales for improving predictions

Integrating new atmospheric model features developed during phase 2

New cloud and convection features:

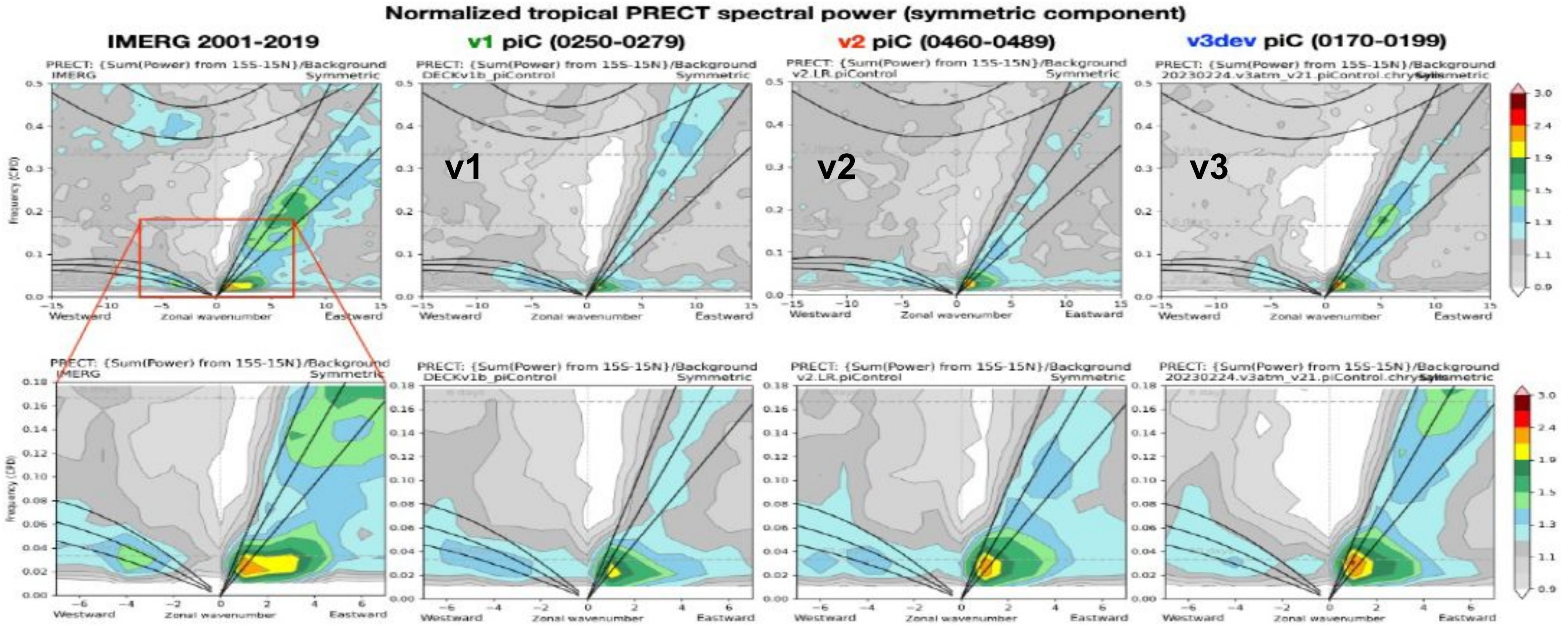
- Predicted Particle Properties (P3) for stratiform clouds → To improve representations of ice particle evolution and inclusion of rimed particles.
- Sophisticated cloud microphysics in Zhang-MacFarlane (ZM) deep convection scheme → Allows aerosols to impact convective processes (through microphysics)
- Multiscale Coherent Structures Parameterization (MCSP) → Represents the effects of organized mesoscale convective systems
- ZM mass flux adjustment to large-scale dynamics → Incorporates the influence of large-scale circulation on deep convection

New chemistry and aerosol features:

- UCI chemistry with 32 transported species + SOA, dust, stratospheric aerosols, nitrate aerosols

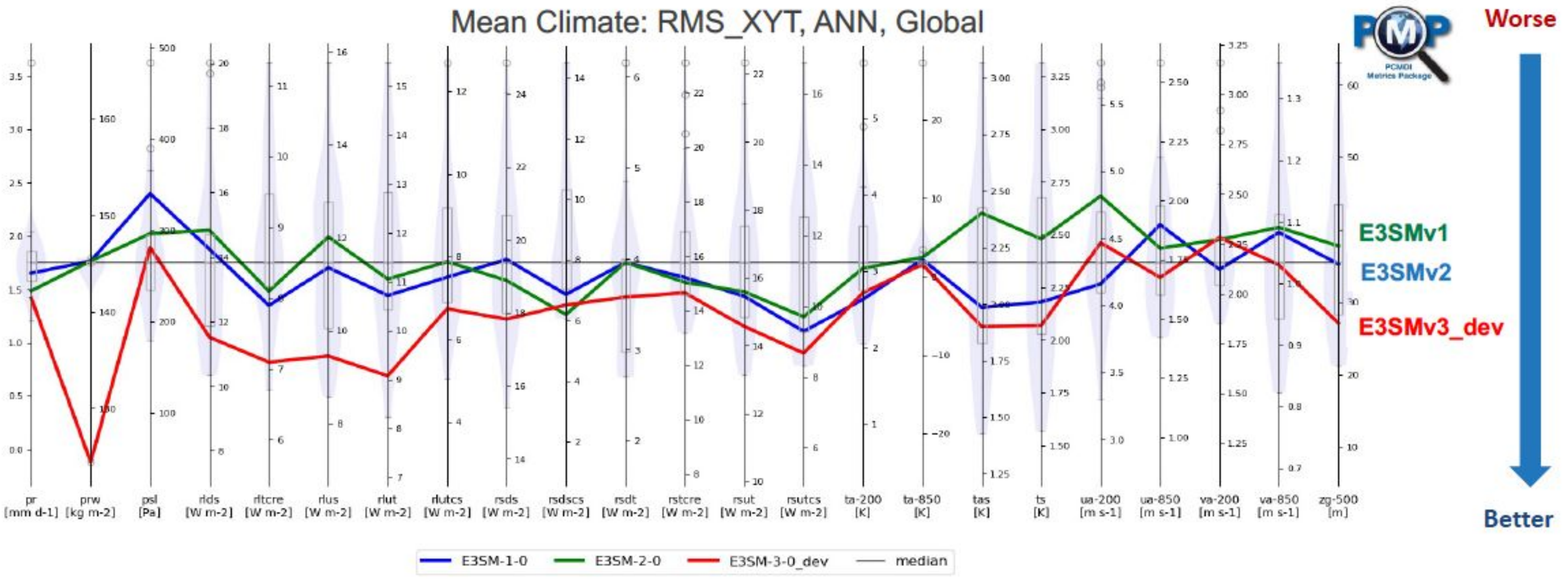
(Shaocheng Xie et al.)

V3_dev shows improvements in many aspects



(Shaocheng Xie et al.)

V3_dev shows improvements in many aspects



(Jiwoo Lee, Jill Zhang, and Xue Zheng)

Wheeler-Kiladis Diagram

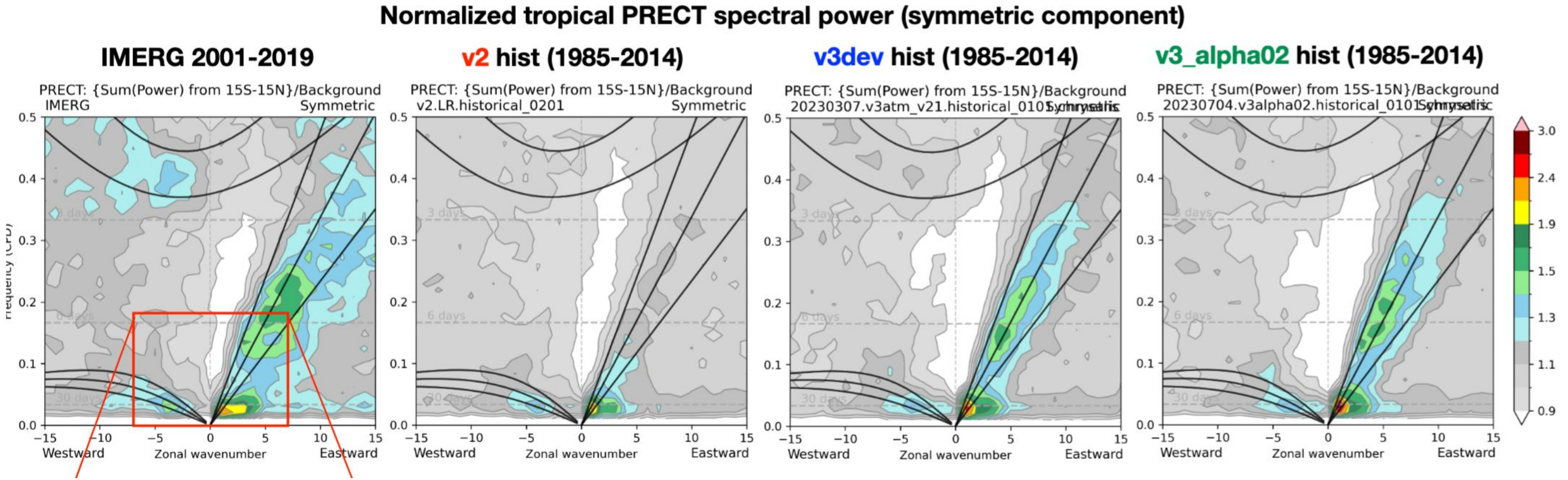


Figure from Jim Benedict

E3SMv3_dev Tropical Variability

Lag Correlation: PRECT & U850 Using East Indian PRECT Index

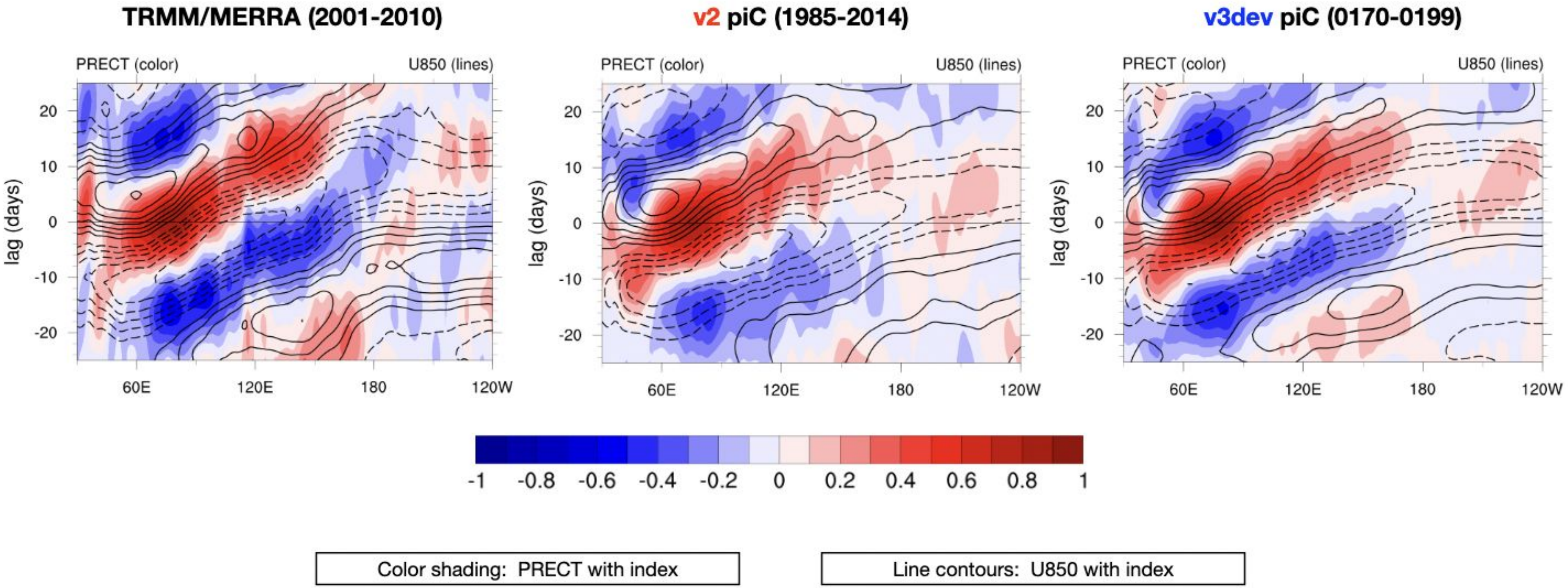


Figure from Jim Benedict