Applications of Uncertainty Quantification (UQ) in Regional and Global Climate Modeling: Parametric Sensitivity Analysis and Calibration/Autotunng

Yun Qian

Pacific Northwest National Laboratory, Richland, WA, USA

Workshop on Uncertainty Quantification in Climate Modeling and Projection

The Abdus Salam International Center for Theoretical Physics, Trieste, Italy

17 July, 2015

Pacific Northwest
NATIONAL LABORATORY

Outline

1. **Parametric sensitivity study based on CAM5 (forward modeling)**

- Precipitation (including extremes and diurnal cycle)
- Structure error
- Short ensemble simulations strategy and process-level calibration
- Aerosol effects

2. **Calibration and autotuning (inverse modeling)**

- Global model CAM5 with a focus on convective precipitation ratio
- Regional model WRF and RegCM3 with a focus on convection parameterization schemes

1. CAMS Sensitivity Analysis

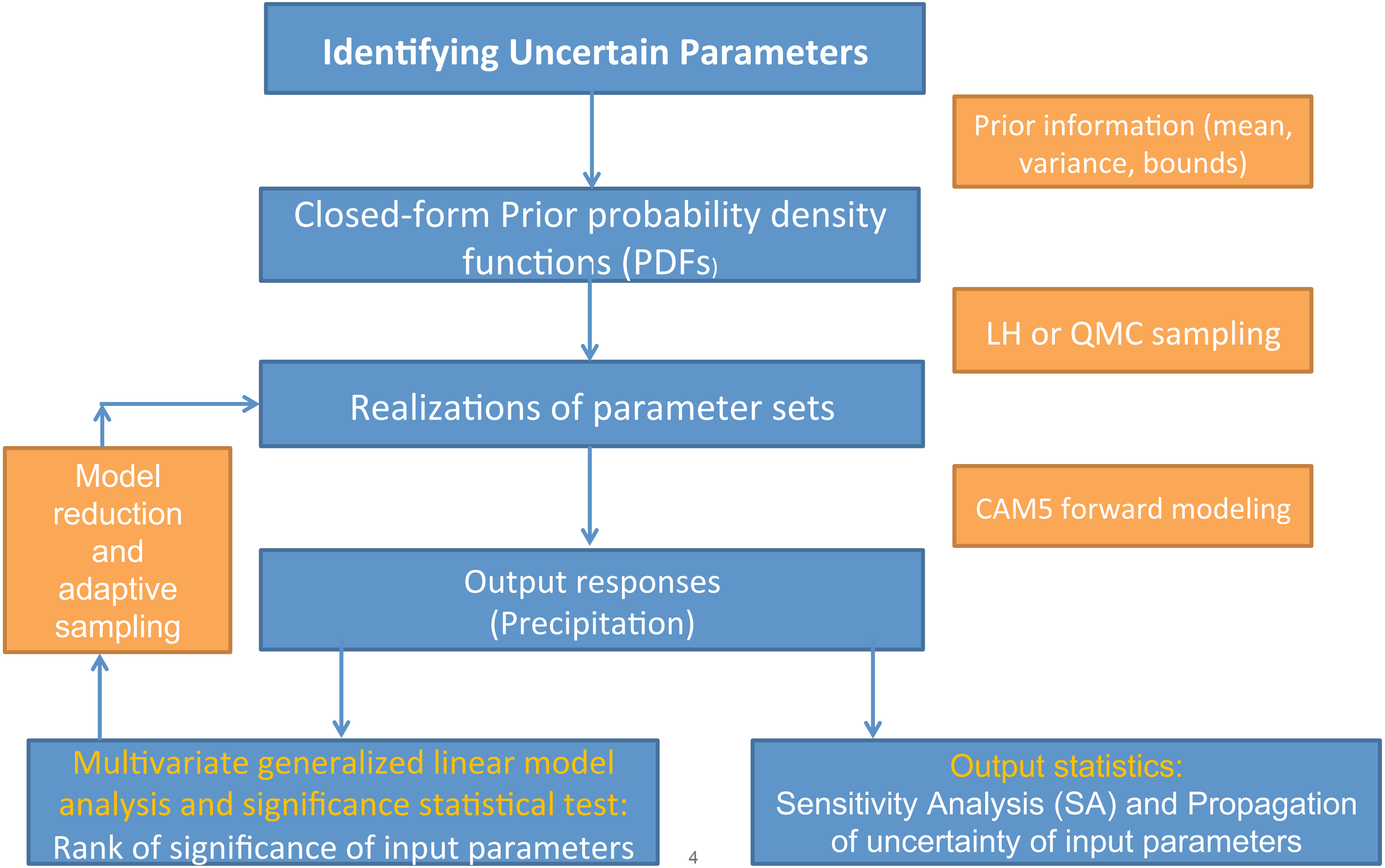
Science Questions:

1. Are there parameters that can dramatically influence the precipitation in CAM5?
2. If yes, how does the parametric sensitivity vary with scale/region/season?
3. Does the parametric sensitivity change with the sampling method or accompanying parameters?
4. What is the relative contribution from individual parameters versus their interactions?

Answering these questions could help:

1. better understand the CAM5 model behavior and physical processes associated with the parameter uncertainties and external forcings
2. guide hand-on model tuning
3. calibrate and optimize model performance

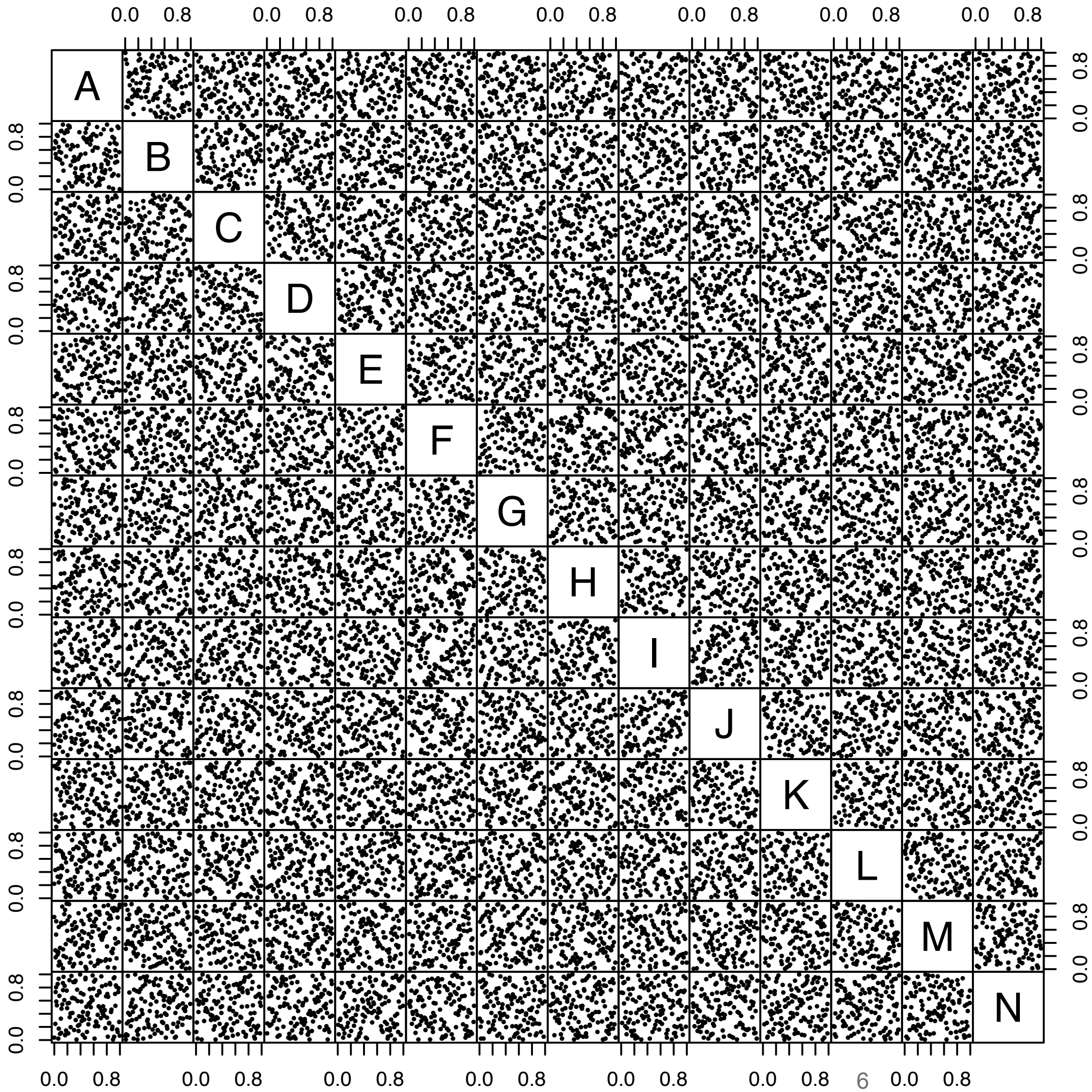
CAM5 Sensitivity Analysis (SA)



CESM/CAM5 Uncertain Parameters of Interest (C-Ensemble)

#	Parameter Name	Range			Description	Namelist Prefix	File Name (.F90)
		Low	Default	High			
1	**rhminh**	0.65	0.80	0.85	Threshold relative humidity for stratiform high clouds	cldfrc_	cloud_fraction
2	**rhminl**	0.80	0.8875	0.99	Threshold relative humidity for stratiform low clouds	cldfrc_	cloud_fraction
3	**alfa**	0.05	0.10	0.60	Maximum cloud downdraft mass flux fraction	zmconv_	zm_conv
4	**c0_lnd**	1.0e-3	0.0059	0.01	*Deep* convection precipitation efficiency over land	zmconv_	zm_conv
5	**c0_ocn**	1.0e-3	0.045	0.1	*Deep* convection precipitation efficiency over ocean	zmconv_	zm_conv
6	**dmpdz**	-2.0e-3	-1.0e-3	-0.2e-3	Parcel fractional mass entrainment rate	zmconv_	zm_conv
7	**ke**	0.5e-6	1.0e-6	10.0e-6	Evaporation efficiency of precipitation	zmconv_	zm_conv
8	**tau**	1800.0	3600.0	28800.0	Time scale for consumption rate deep CAPE	zmconv_	zm_conv
9	**ai**	350.0	700.0	1400.0	Fall speed parameter for cloud ice	no nml	cldwat2m_micro
10	**as**	5.86	11.72	23.44	Fall speed parameter for snow	no nml	cldwat2m_micro
11	**cdnl**	0.0	0.0	1.0e+7	Lower bound on droplet number	no nml	cldwat2m_micro
12	**dcs**	100e-6	400e-6	500e-6	Autoconversion size threshold for ice to snow	no nml	cldwat2m_micro
13	**eii**	0.001	0.1	1.0	Collection efficiency aggregation ice	no nml	cldwat2m_micro
14	**qcvar**	0.5	2.0	5.0	Inverse relative variance of sub-grid cloud water	no nml	cldwat2m_micro
15	**a2l**	10.0	30.0	50.0	Moist entrainment enhancement parameter	no nml	eddy_diff
16	**criqc**	0.5	0.7	1.5	Maximum updraft condensate	nml/add	uwshcu
17	**kevp**	1.0e-6	2.0e-6	20.0e-6	Evaporative efficiency	nml/add	uwshcu
18	**rkm**	8.0	14.0	16.0	Updraft lateral mixing efficiency	nml/add	uwshcu
19	**rpen**	1.0	5.0	10.0	Penetrative updraft entrainment efficiency	uwshcu_	uwshcu
20	**e_dust**	0.21	0.43	0.86	Dust emission tuning factor	aerosol_	aerosol_intr^
21	**wsubimax**	0.1	0.2	1.0	Maximum subgrid vertical velocity for ice nucl	x	microp_aero
22	**Wsubmin**	0.0	0.2	1.0	Minimum subgrid vertical velocity for liquid nucl	x	microp_aero

PPE C-Ensemble (LHP)



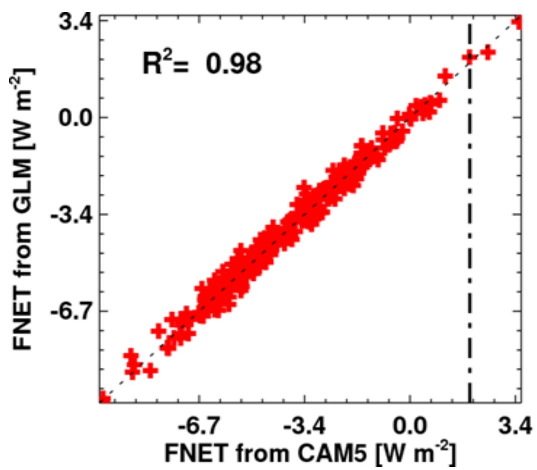
- LLNL: C-Ensemble
- Latin Hypercube
- 22 parameters
- 1100 sample sets (forward simulations)
- Each simulation: 5-yr
- Each parameters is sampled 1100 times

Surrogate Model: Generalized Linear Model (GLM)

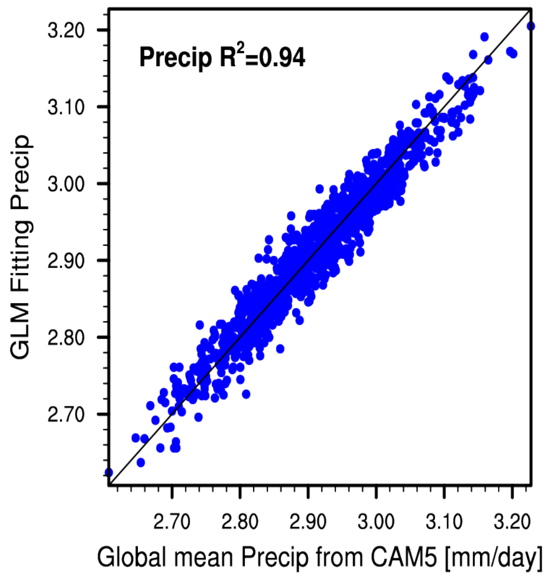
Pacific Northwest
NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

GLM is a flexible generalization of ordinary linear regression that allows for response variables that have other than a normal distribution

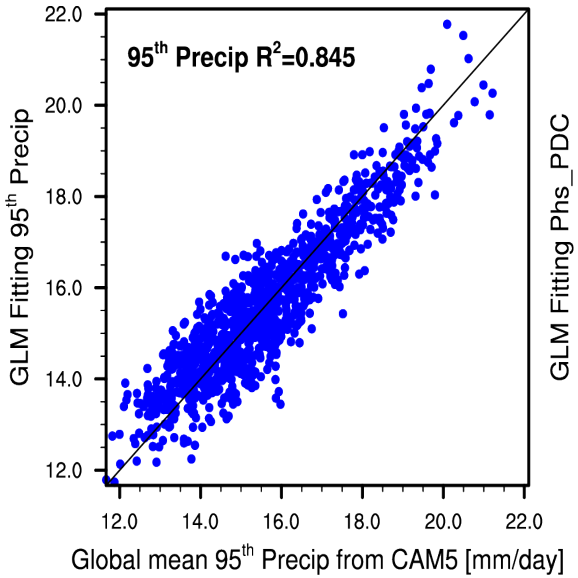
$$Y_i = \beta_0 + \sum_{j=1}^{10} \beta_j * p_{i,j} + \varepsilon_i, \qquad \varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$$



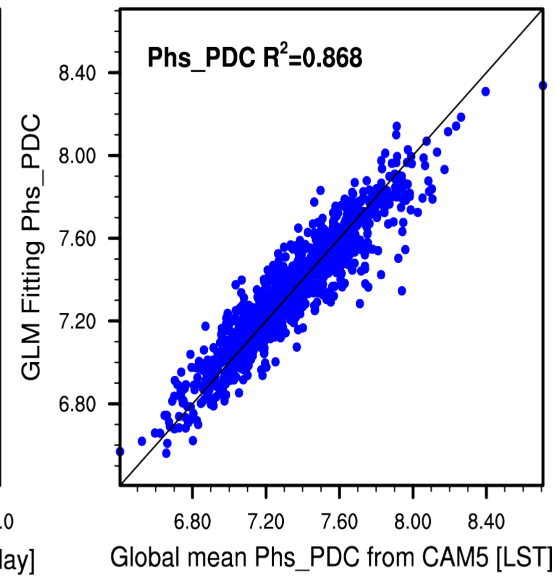
GLM-fitted global precipitation versus the CAM5 simulations



Mean

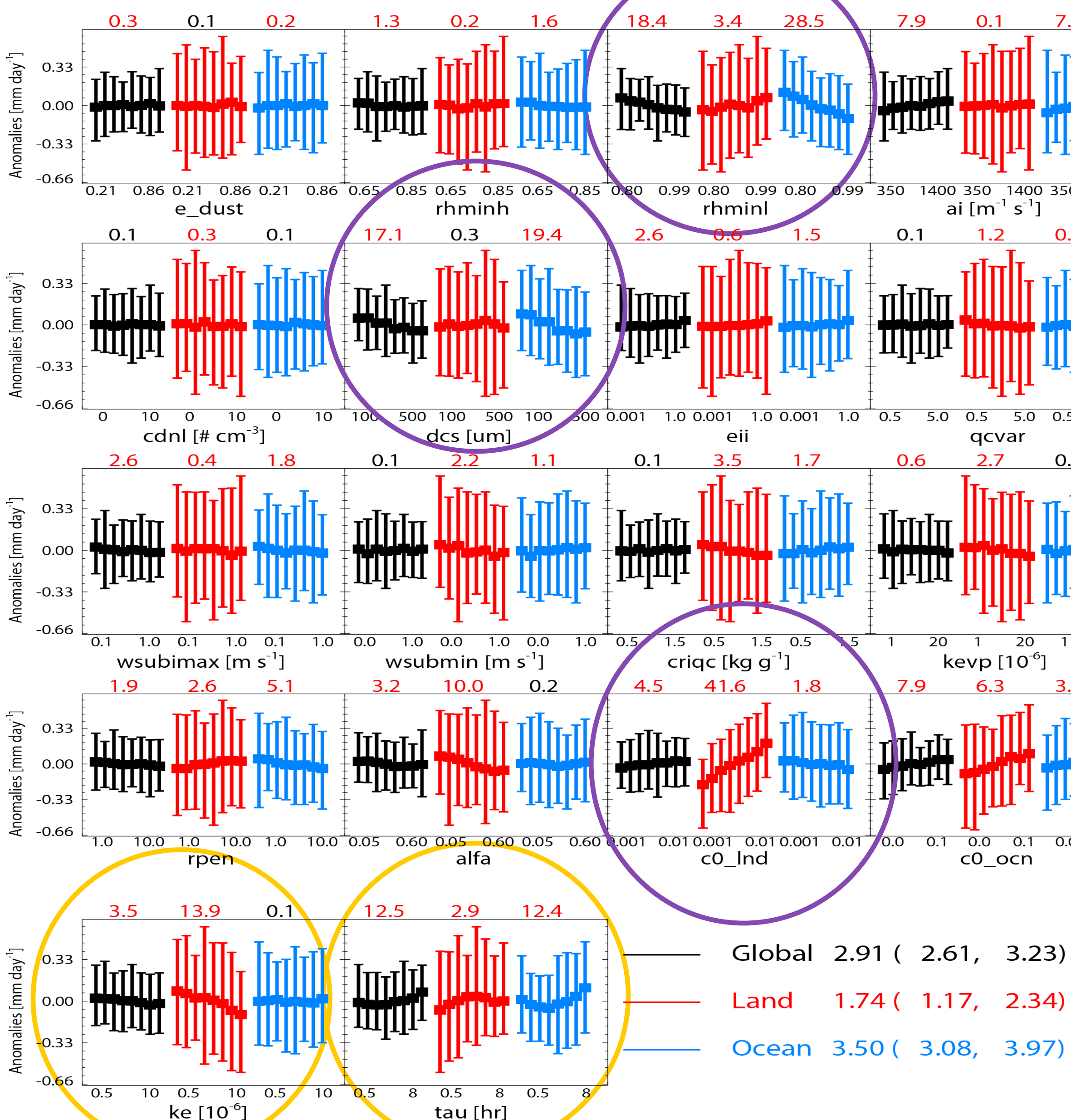


95th

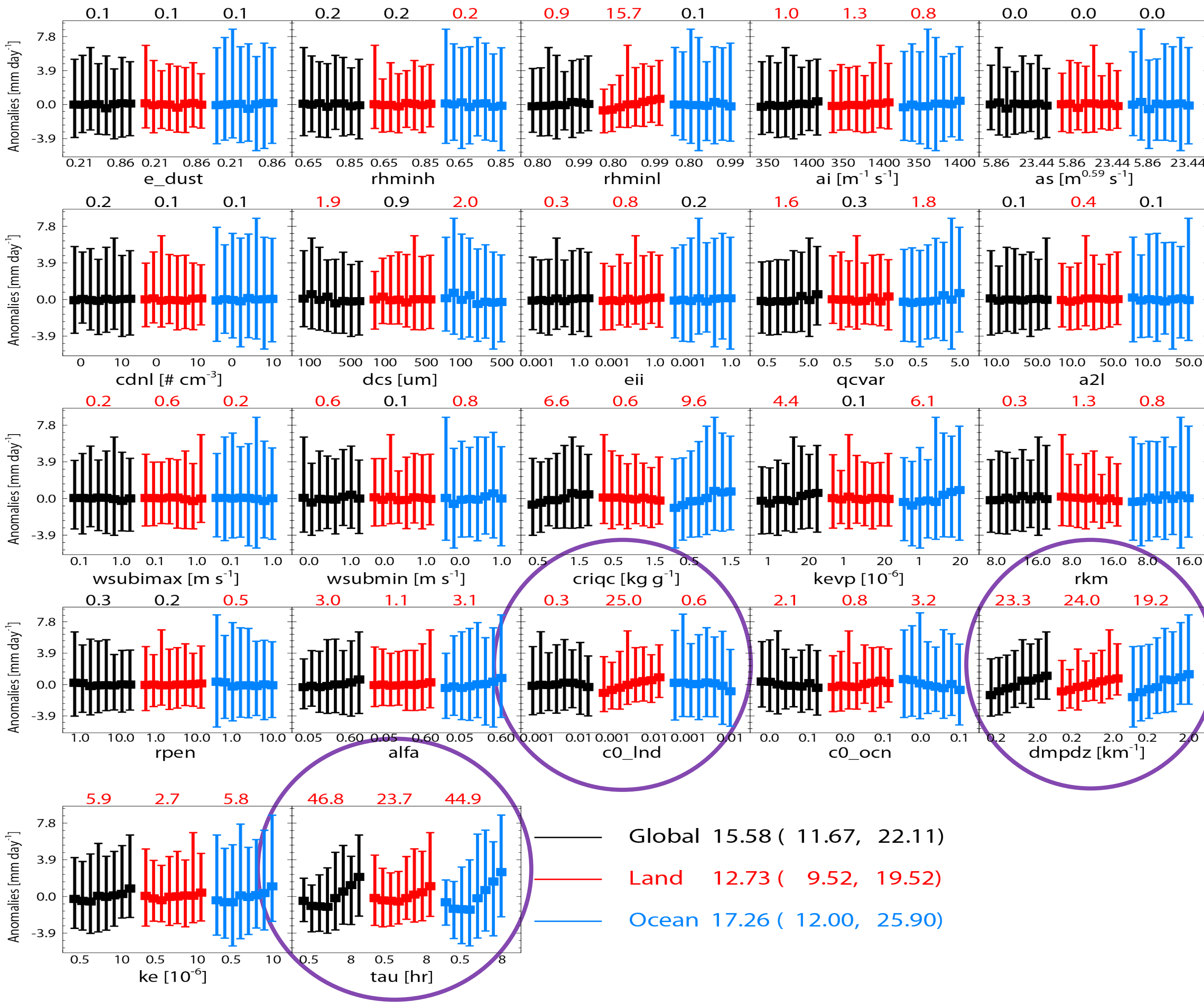


Phase

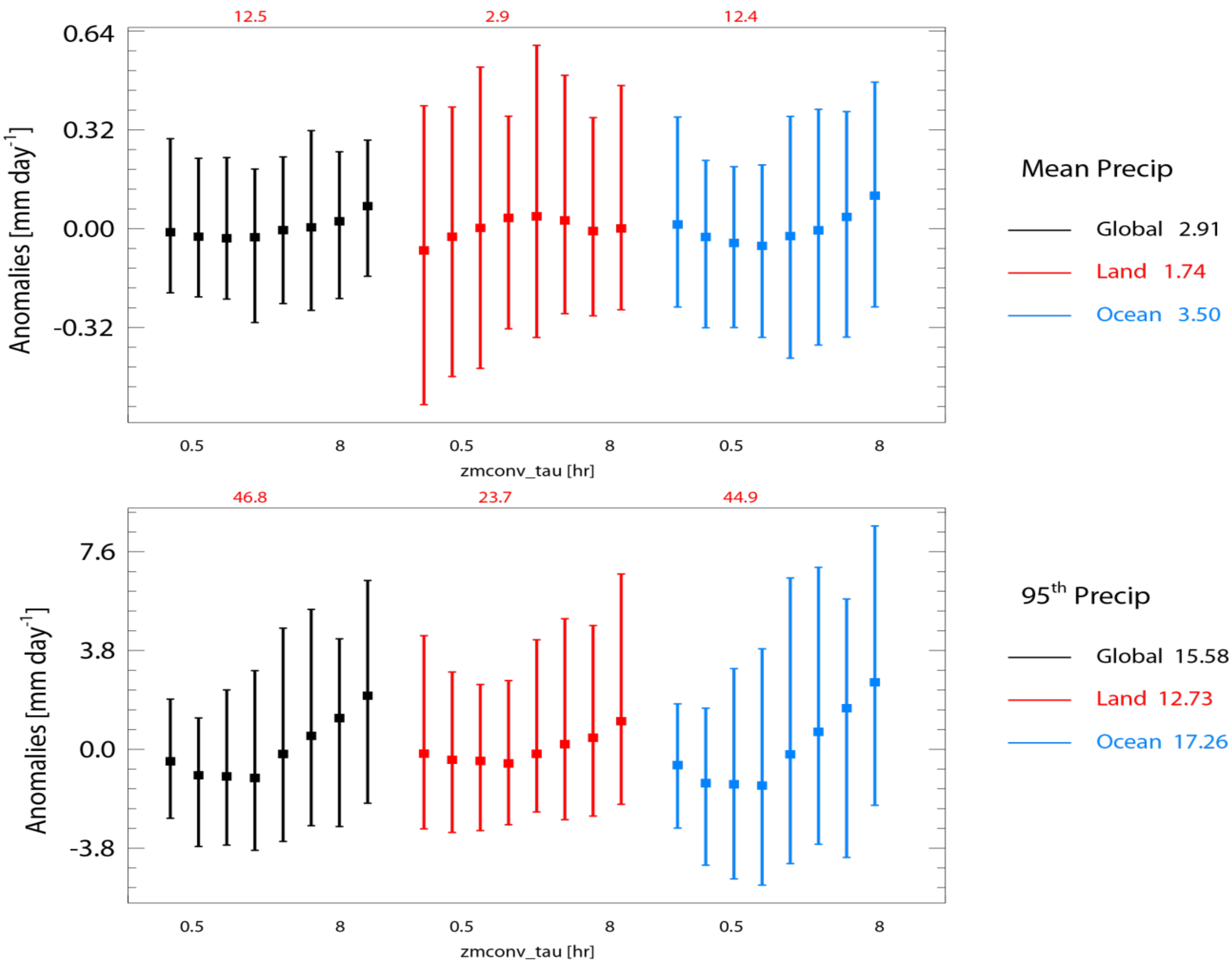
Sensitivity of mean precipitation to C-Ensemble parameters



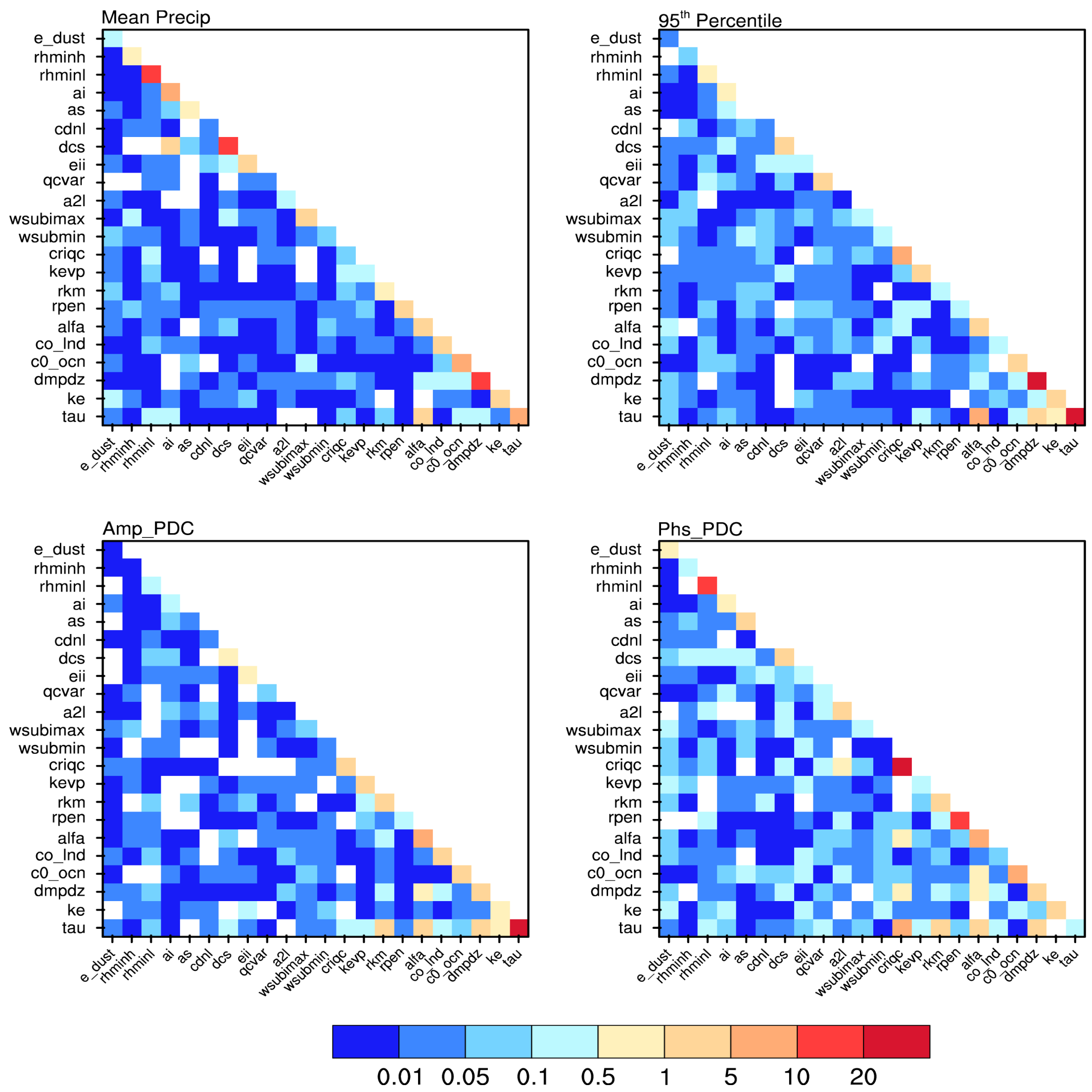
Sensitivity of 95th precipitation to C-Ensemble parameters



CAPE consumption time scale (tau)



Relative contributions (%) of individual parameter and their interactions (C-Ensemble)

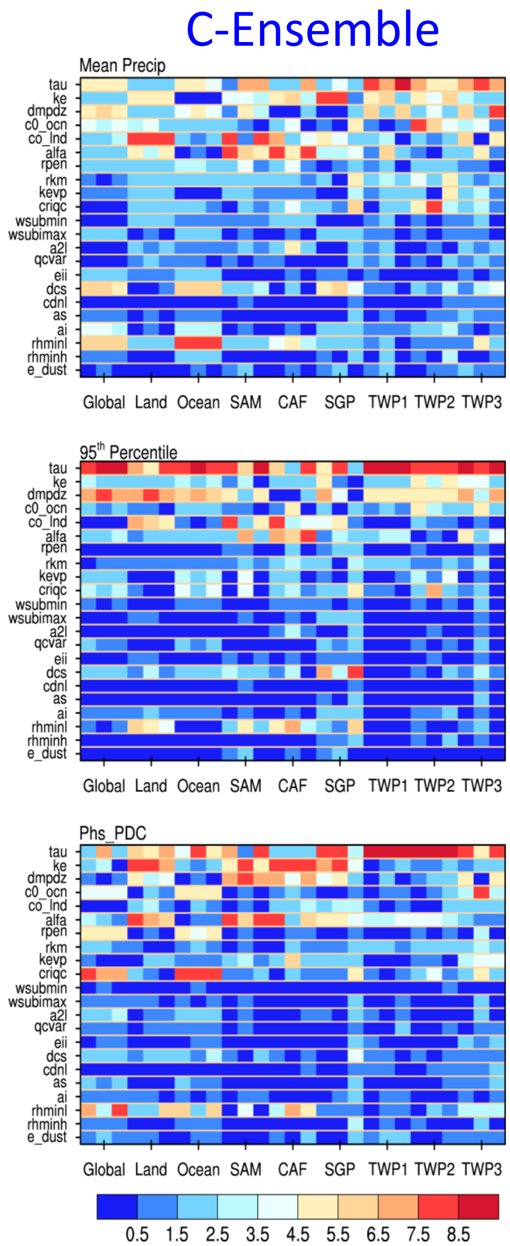


Sensitivity of each parameter at different region/scale/season

Mean precipitation

95th precipitation

Phase of PDC



Summary (Qian et al, JAMES, 2015)

Pacific Northwest NATIONAL LABORATORY

Proudly Operated by Battelle Since 1965

- We investigated the sensitivity of precipitation characteristics to dozens of uncertain parameters mainly related to cloud processes in the CAM5.
- Most sensitive parameters to

  Mean Precip: *c0_lnd, rhminl, dcs, tau, dmpdz,* and *ke*

  Extreme Precip: <u>*tau*</u> *(~50% total variance), c0_lnd, dmpdz*

  Phase of Diurnal Cycle: *ke, alfa* and *tau*

  *Precipitation not monotonically respond to *tau* (a turning point ~ 1.75 hours)
- The influence of individual parameters does not depend on the sampling approach applied or concomitant parameters selected.
- The total variance for precipitation is primarily contributed by the individual parameters (75-90% in total), and their interactions contribute to the rest of total variance explained.

Structure Error

1100 C-Ensemble Variance

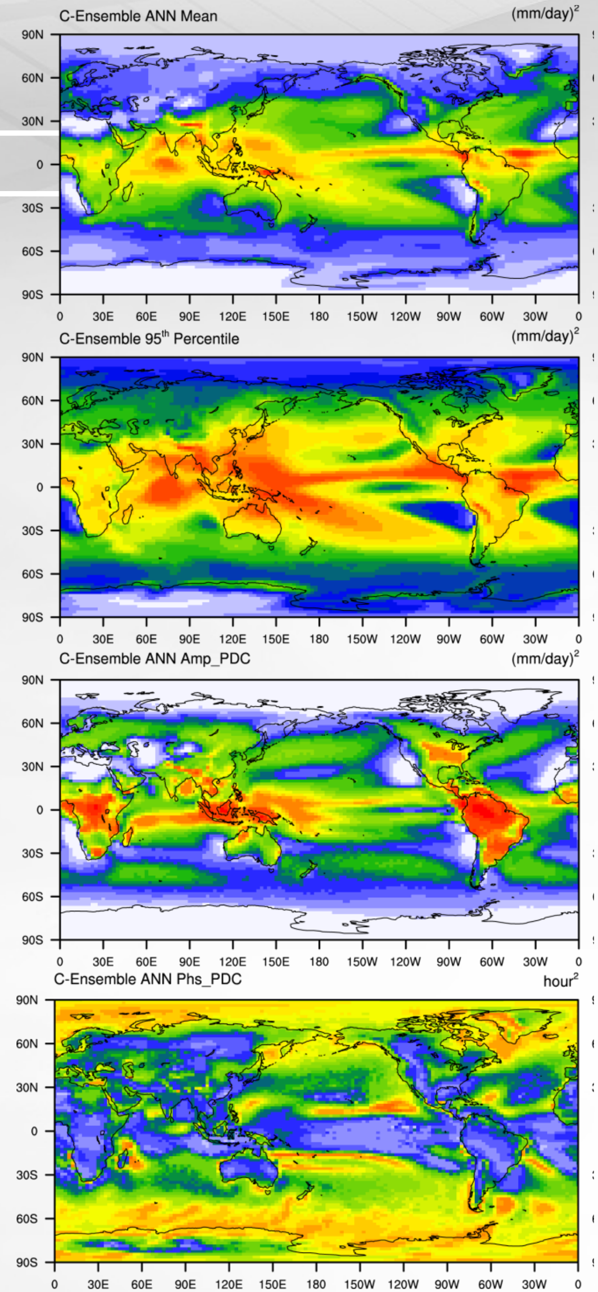
Pacific Northwest
NATIONAL LABORATORY
Proudly Operated by Battelle *Since 1965*

Mean Precip

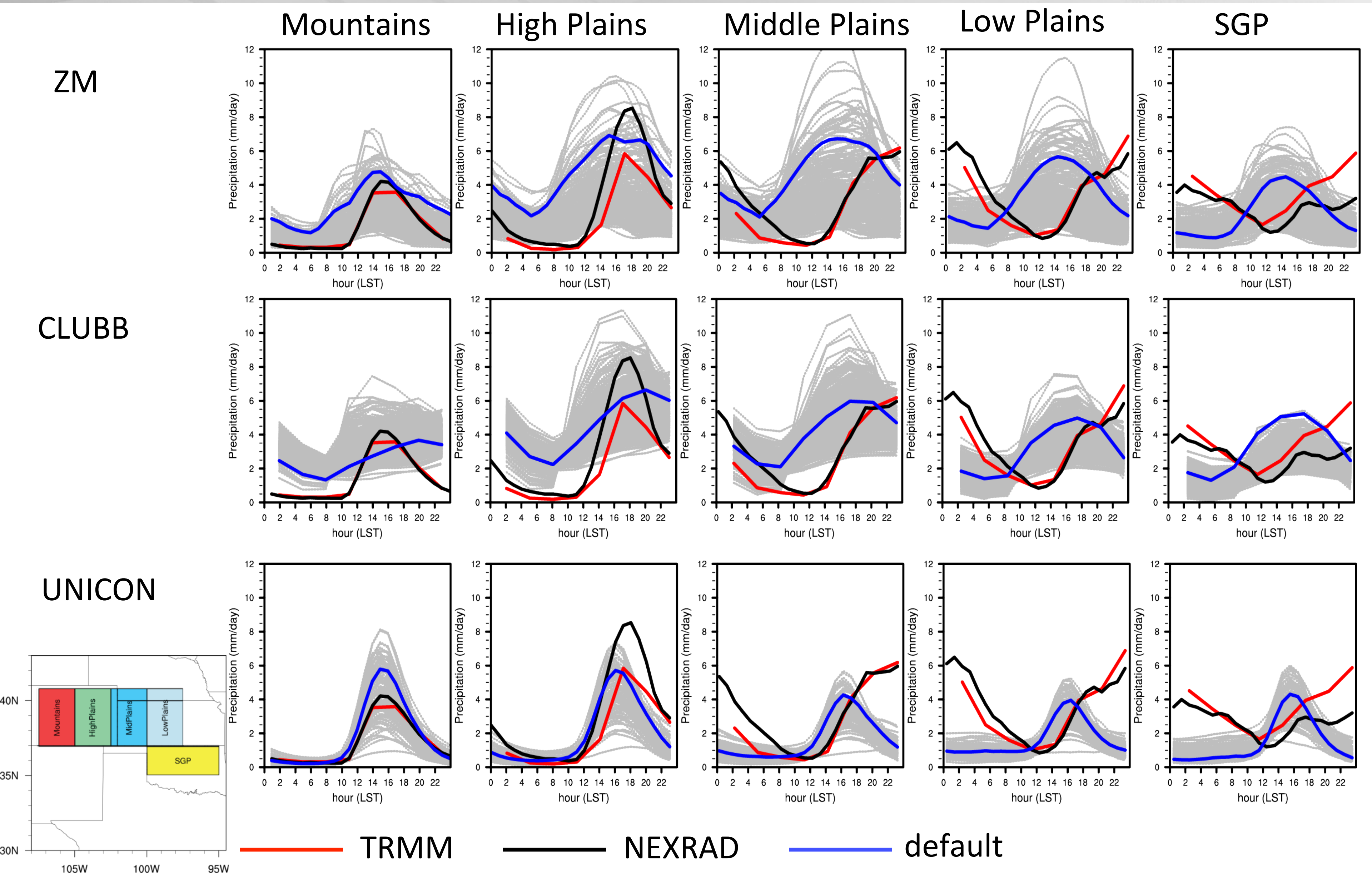
95th percentile

Amplitude of DC

Phase of DC

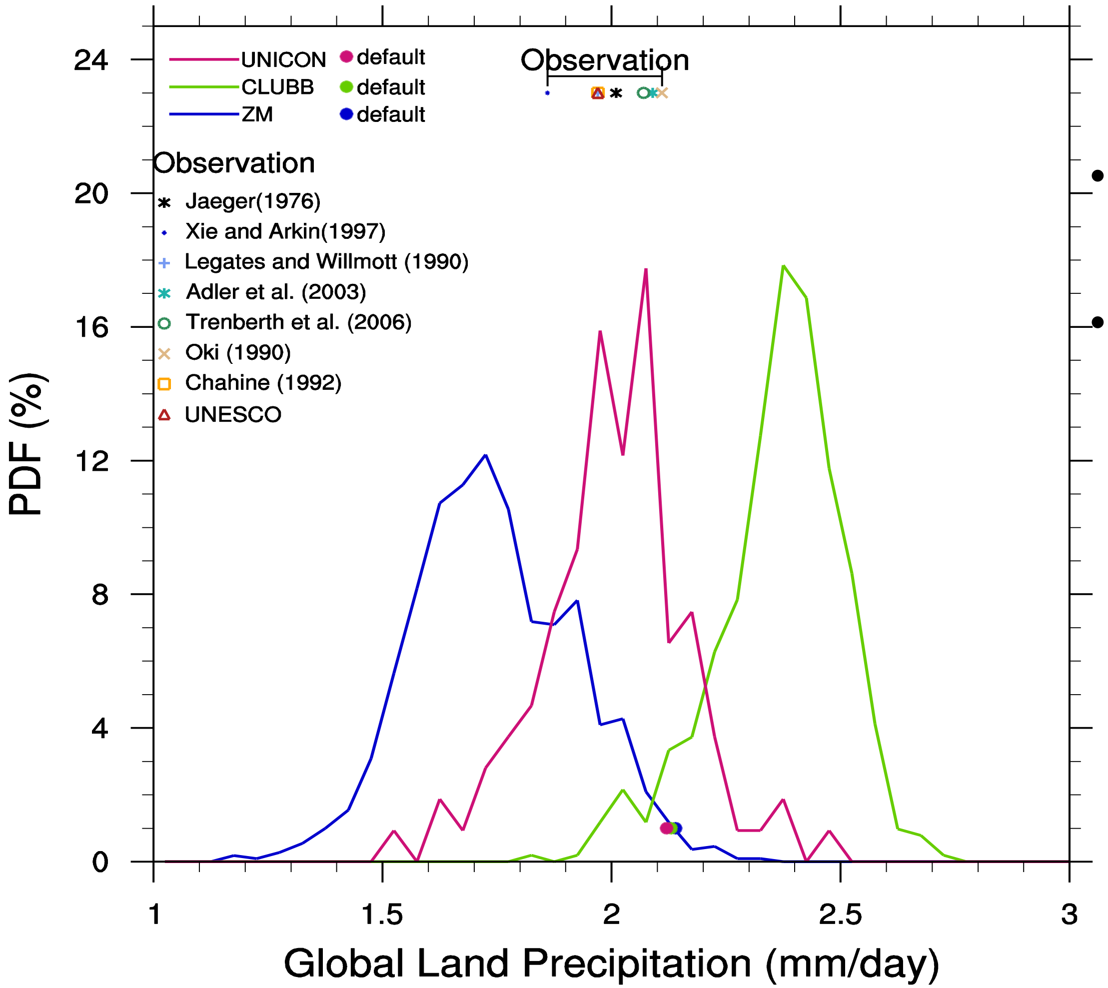


Diurnal cycle and MCS propagation over central US



Land mean precipitation

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965



- Defaults are very similar (fine-tuned?)
- BUT the PDFs and means/ medians are clearly different!

Short ensemble simulations strategy and process-level calibration

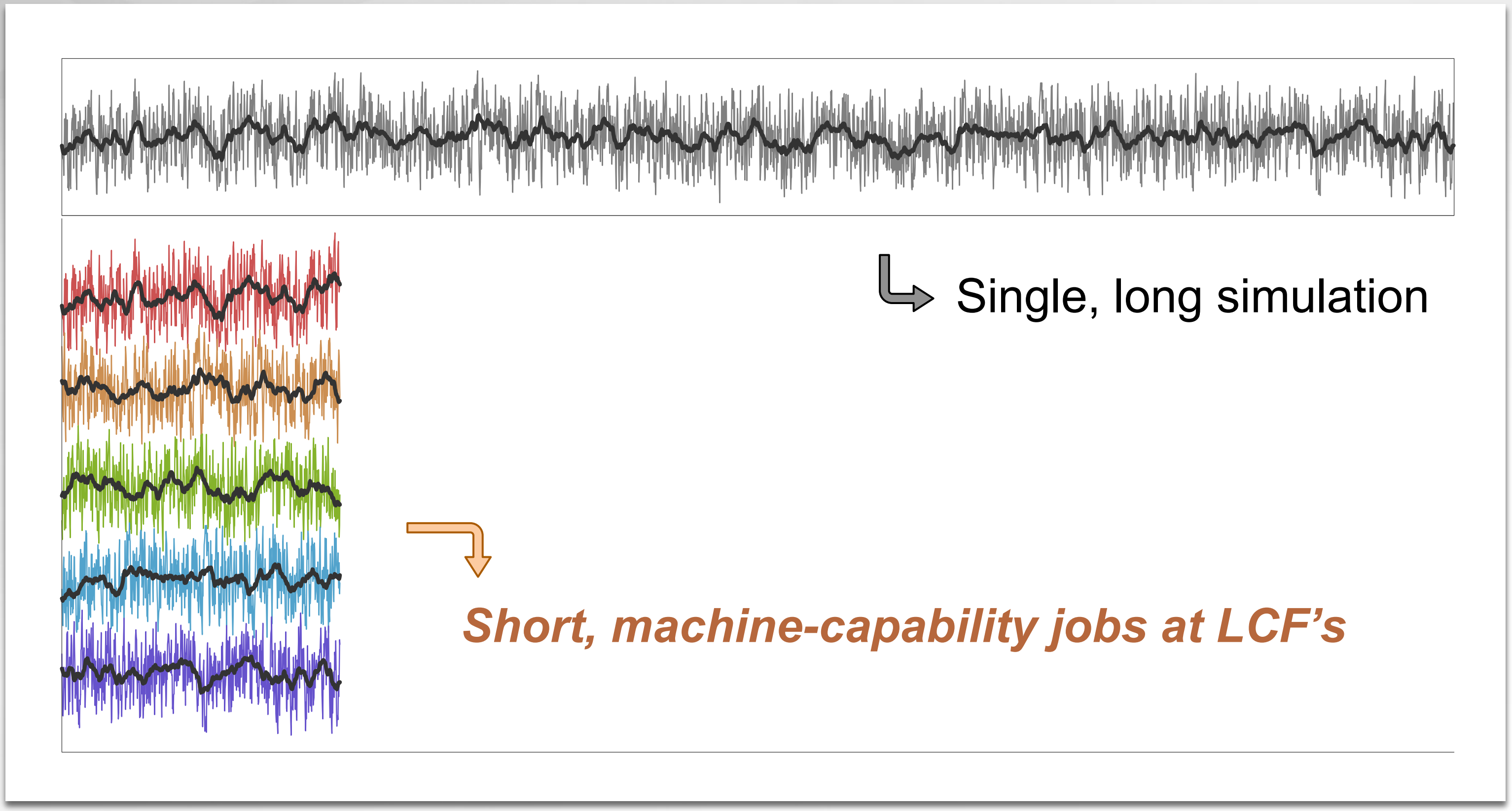
Short (Few-day) Simulations for Efficient Model Evaluation, Tuning and Calibration

Process-level (e.g. cloud, convection) calibration

ACME Needs New, Efficient Strategies for Model Evaluation and Tuning

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

- High-resolution, multi-decade simulations are hugely expensive

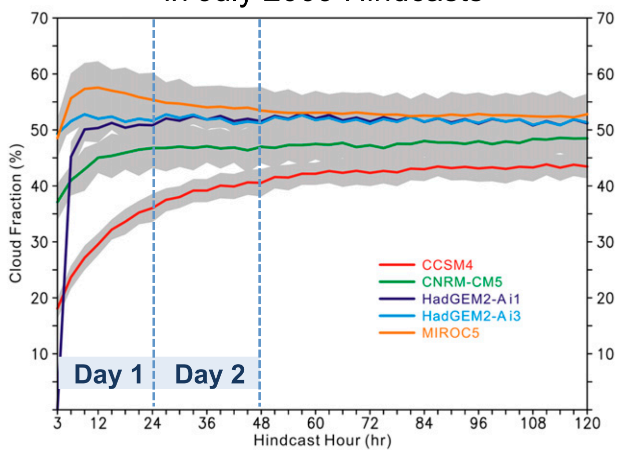


Previous Successes

- Fast processes, especially those related to clouds, are a major source of biases in current climate models

CAPT and Transpose-AMIP

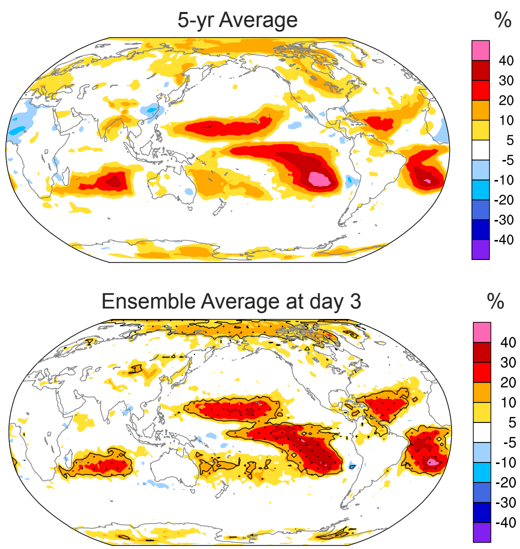
Global Mean Total Cloud Fraction
in July 2009 Hindcasts



Ma et al. (2014)

SciDAC Multiscale

Total Cloud Fraction Difference, CAM5
4-min minus 30-min Time Step

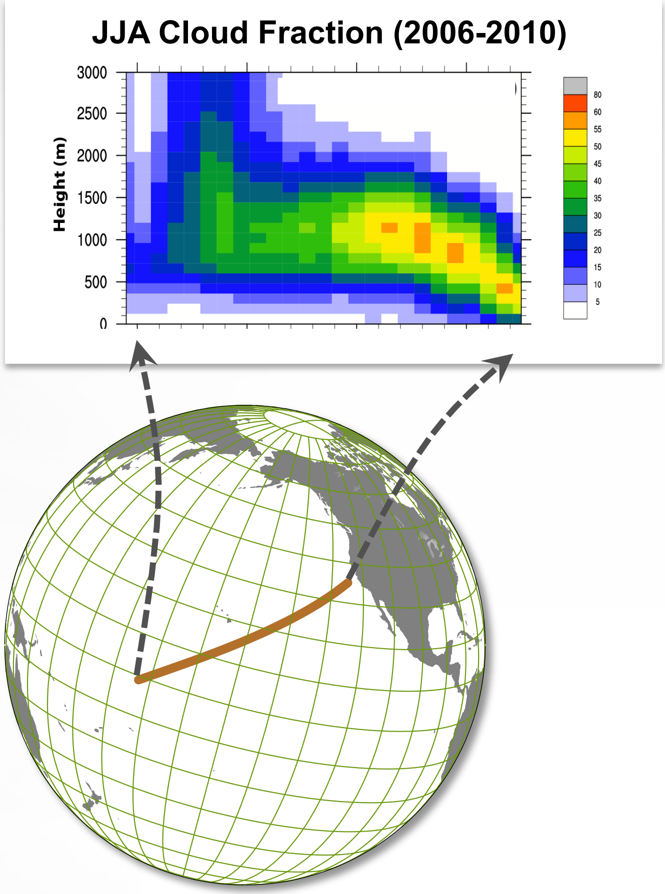


Wan et al. (2014)

Short Simulations Task

- Explore **few-day simulations** for model tuning and sensitivity studies
- **Two-phase investigation**
  - Parametric sensitivity experiments
  - Automatic parameter tuning

- Extensive use of **UQ techniques**
  - Sensitivity analysis
    Qian et al. (2015), Guo et al. (2014, 2015), Zhao et al. (2013)
  - Model calibration and auto-tuning
    Yang et al. (2012, 2013, 2014), Zou et al. (2014)



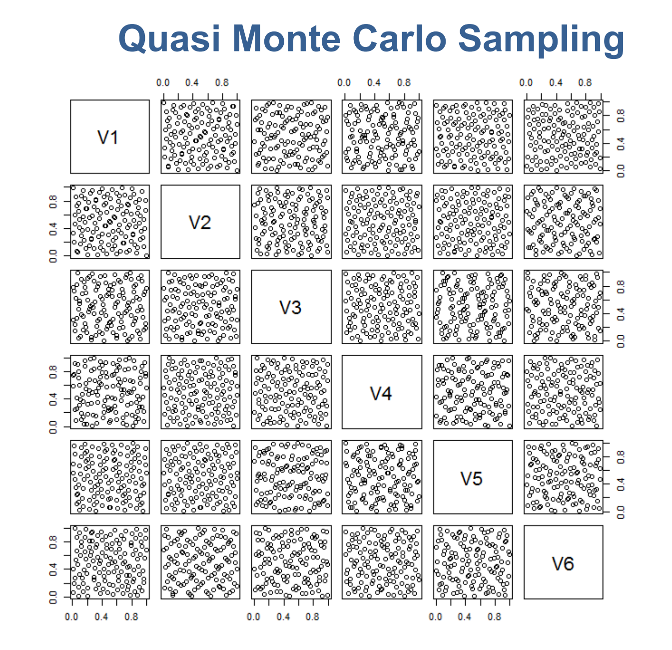
Focus Region:
the GPCI Cross-section

Preliminary results:

▶ A **framework** for short-ensemble-based parametric sensitivity experiments

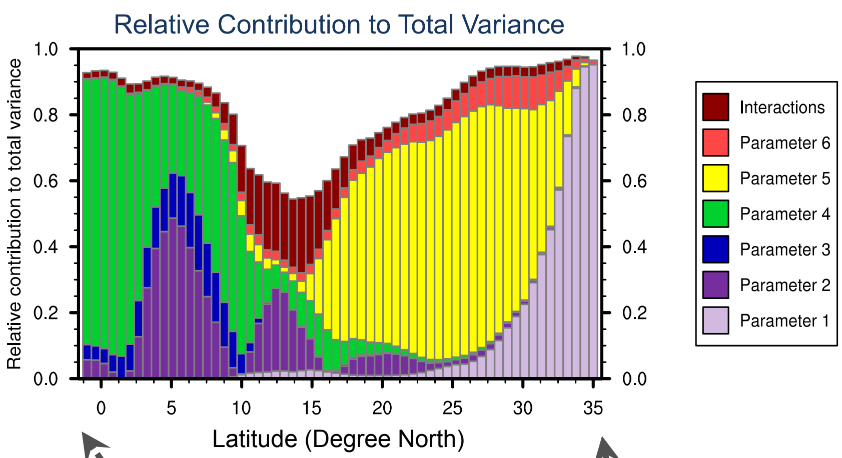
- **31x128 CAPT hindcasts** for July 2008
  - 1 degree resolution (ne30)
  - Using the multi-instance capability for simulation bundling
  - Finished within 3 days(!) on Titan

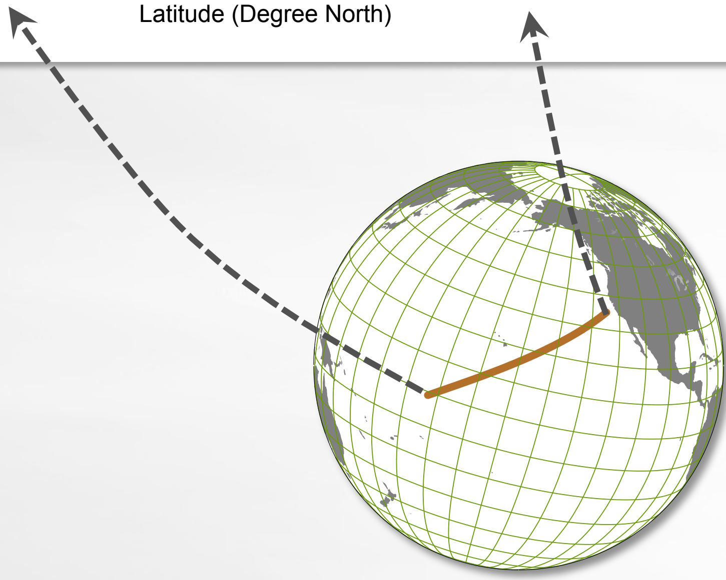
- Parametric **sensitivity analysis**
  - 6 uncertain parameters related to turbulence and shallow convection
  - Quasi Monte Carlo method for sampling parameter space
  - Surrogate model for parametric sensitivity analysis



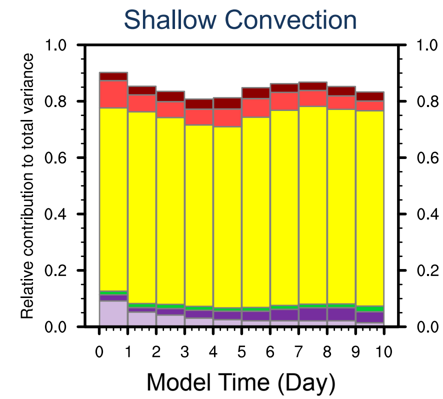
Parametric Sensitivity of Shortwave Cloud Forcing

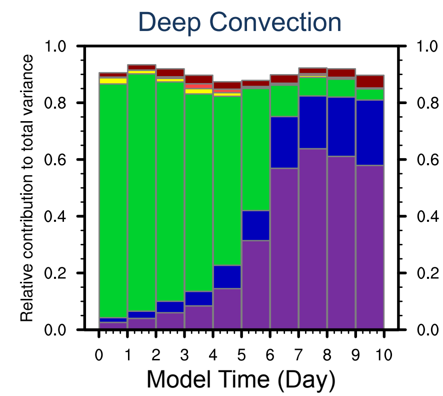
Dependence of Model Sensitivity on Cloud Regime





Time Evolution

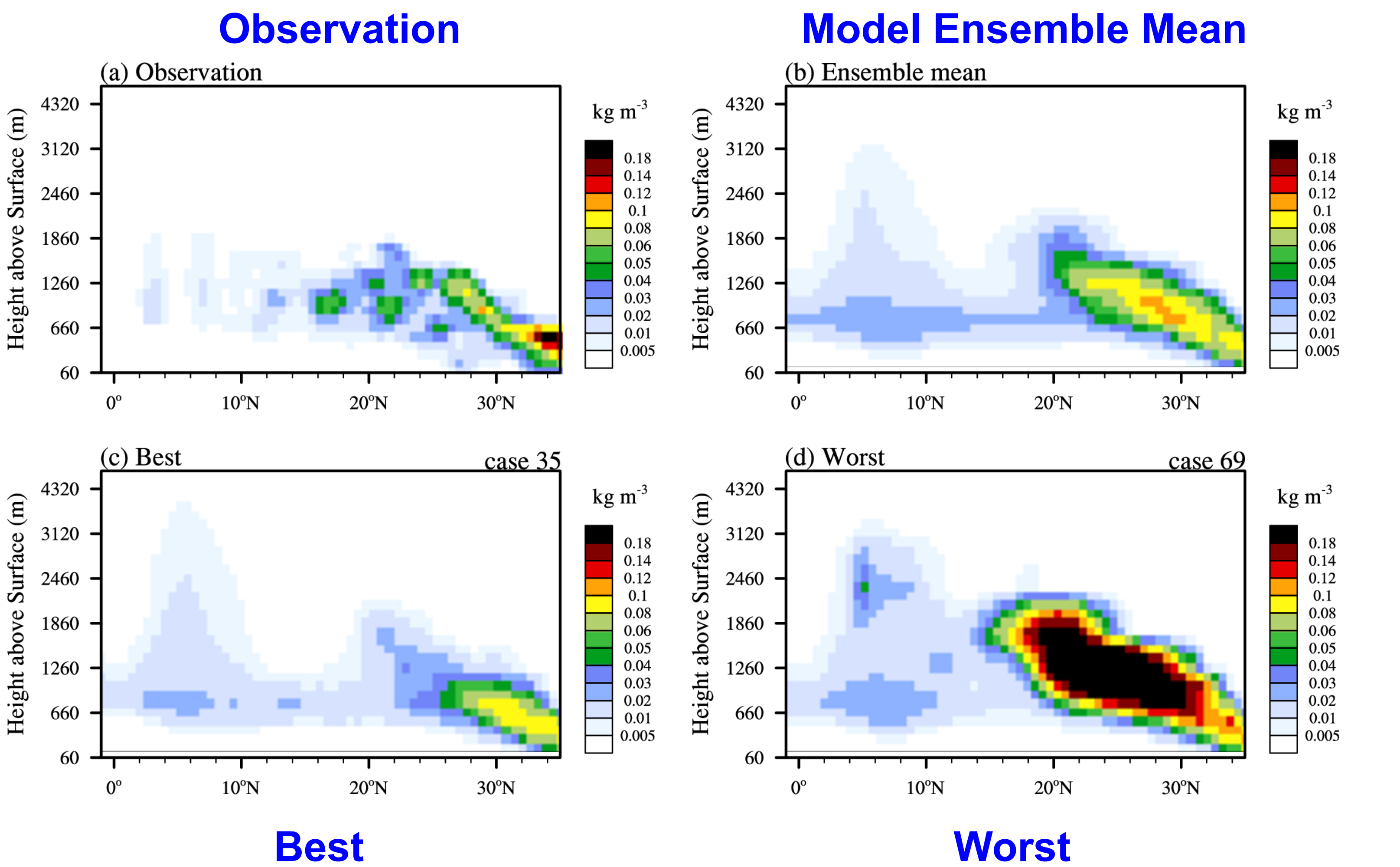




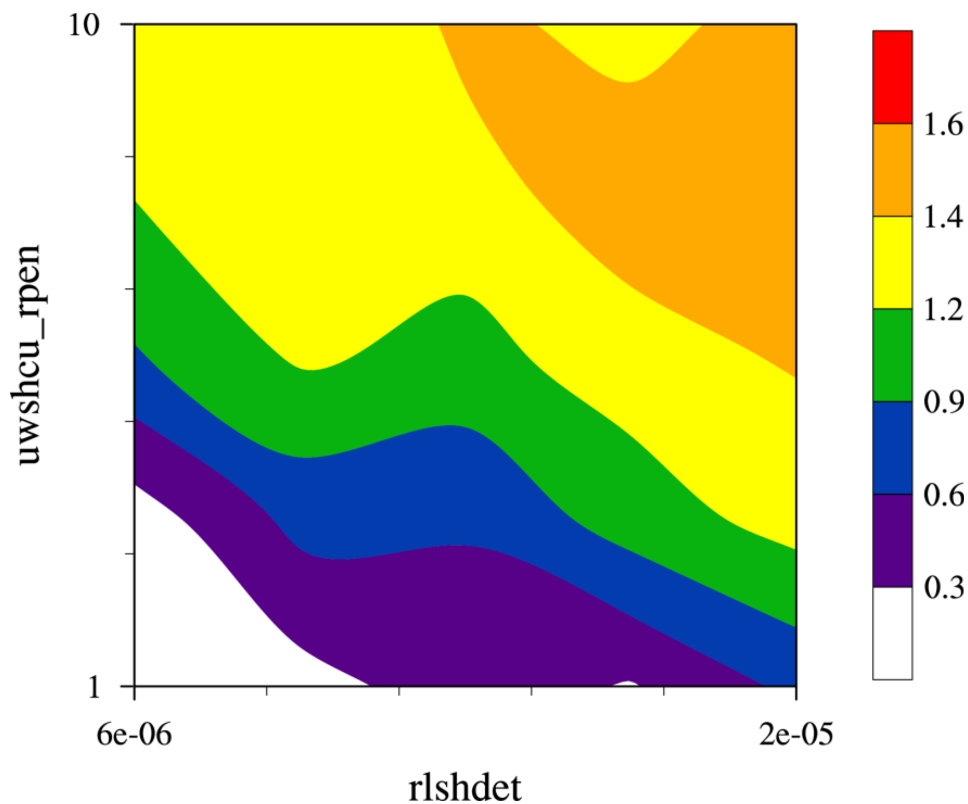
Atmosphere Group, Short Simulations Task Team

Liquid Water Content (LWC) for day 5

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965



Posterior joint 2D marginal distribution
(uwshcu_rpen vs. r1shdet)



day 5

The joint 2D marginal PDFs are the results of integration across the other 4 dimensions of parameters.

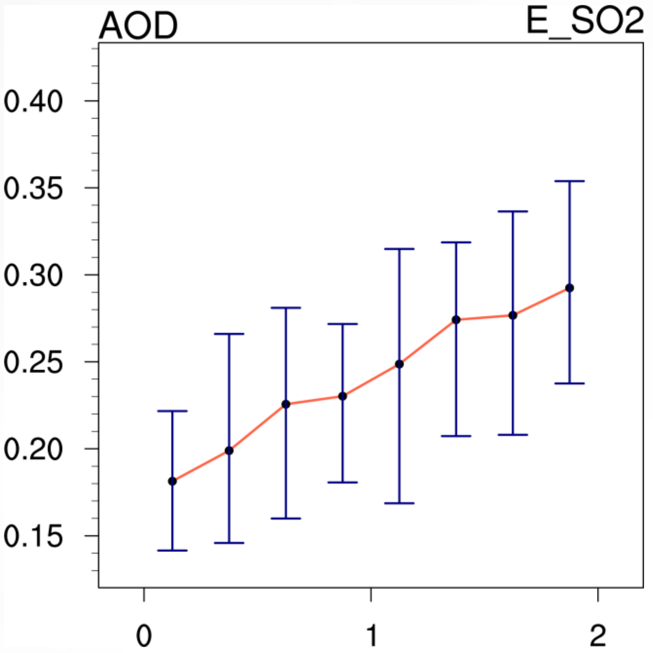
Aerosol Effects

- A New Approach to Modeling Aerosol Effects on East Asian Climate: Parametric Uncertainties Associated with Emissions, Cloud Microphysics and Their Interactions

0 ---------------- 1

1 ---------------- 2

Control **Sensitivity**



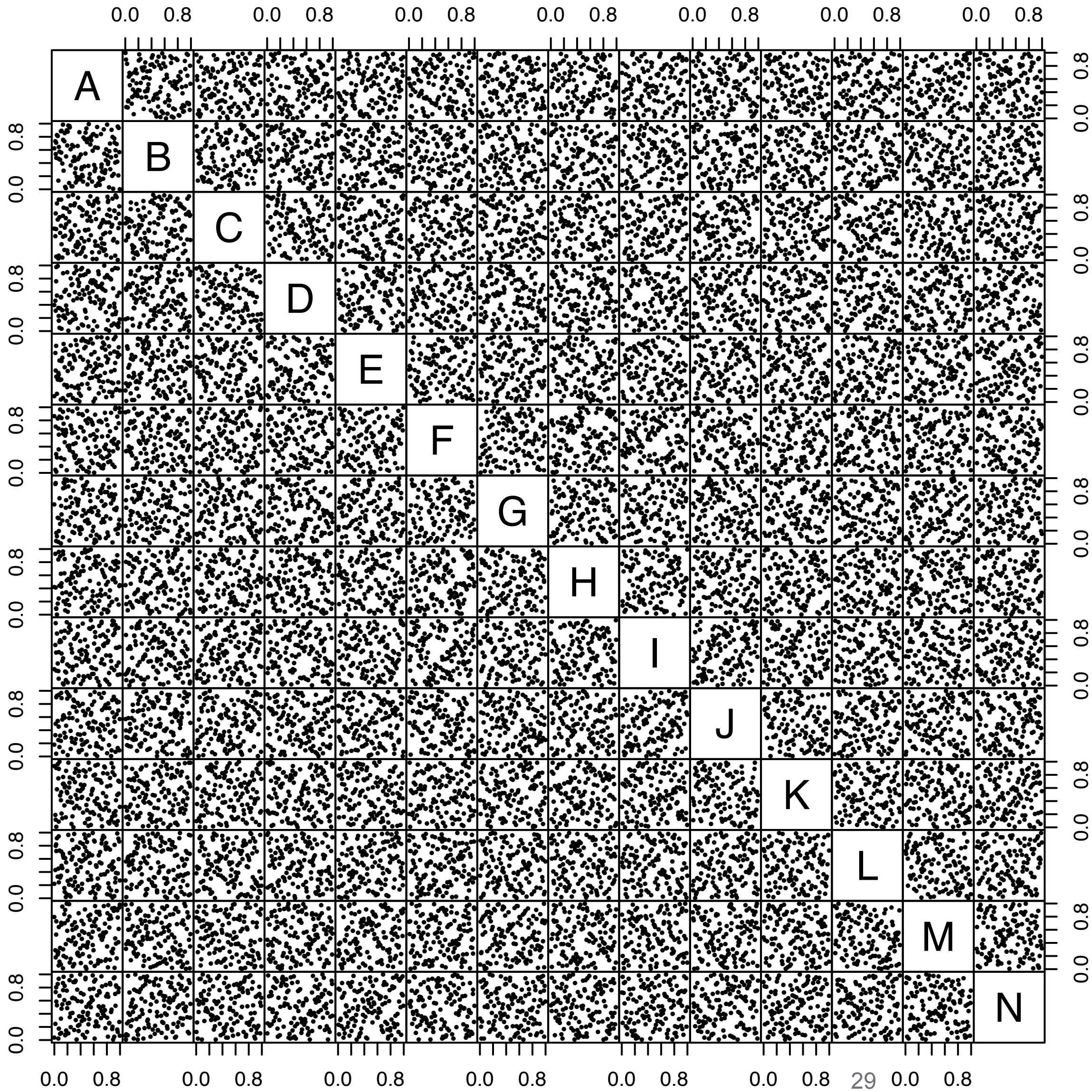
CESM/CAM5 Uncertain Parameters of Interest (A-Ensemble)

#	Parameter Name	Range Low	Range Default	Range High	Description	Namelist Prefix	File Name (.F90)	
1	ai	350.0	700.0	1400.0	Fall speed parameter for cloud ice	cldwatmi_	cldwat2m_micro	M
2	as	5.86	11.72	23.44	Fall speed parameter for snow	cldwatmi_	cldwat2m_micro	M
3	cdnl	0.0	0.0	10.0e+6	Cloud droplet number limiter	cldwatmi_	cldwat2m_micro	LGE
4	dcs	100.0e-6	400.0e-6	500.0e-6	Autoconversion size threshold for ice to snow	cldwatmi_	cldwat2m_micro	M
5	wsubmin	0.0	0.2	1.0	Minimum sub-grid vertical velocity	micropa_	microp_aero	LGE
6	e_dust	0.21	0.35	0.86	Dust emission tuning factor		aerosol_intr	LGE
7	e_sst	0.5	1.0	2.0	Sea salt emission tuning factor		progsseasalt_intr	LGE
8	e_soag	0.5	1.5	2.0	SOA (g) emission scaling factor		emission file	LGE
9	e_acnum	0.3	1.0	5.0	Number emission scaling factor for fossil fuel aerosol		emission file	LGE
10	sol_factic	0.2	0.4	0.8	Solubility factor for the removal of interstitial aerosols in convective clouds		mz_aerosols_intr	LGE
11	sol_facti	0.5	1	1	Solubility factor for cloud-borne aerosols in stratiform clouds		mz_aerosols_intr	LGE
12	ref_dust	0.001	0.005	0.01	Visible imag refractive index for dust		modal_aero_init_data	LGE
13	e_so2	0	1	2	emission tuning factor for SO2			
14	e_bc	0	1	3	emission tuning factor for BC			
15	e_pom	0	1	3	emission tuning factor for POM		modal_aero_init_data	LGE
16	e_so4f	0	0.025	0.05	emission tuning factor for sulfate		modal_aero_init_data	LGE

PPE A-Ensemble (QMC)

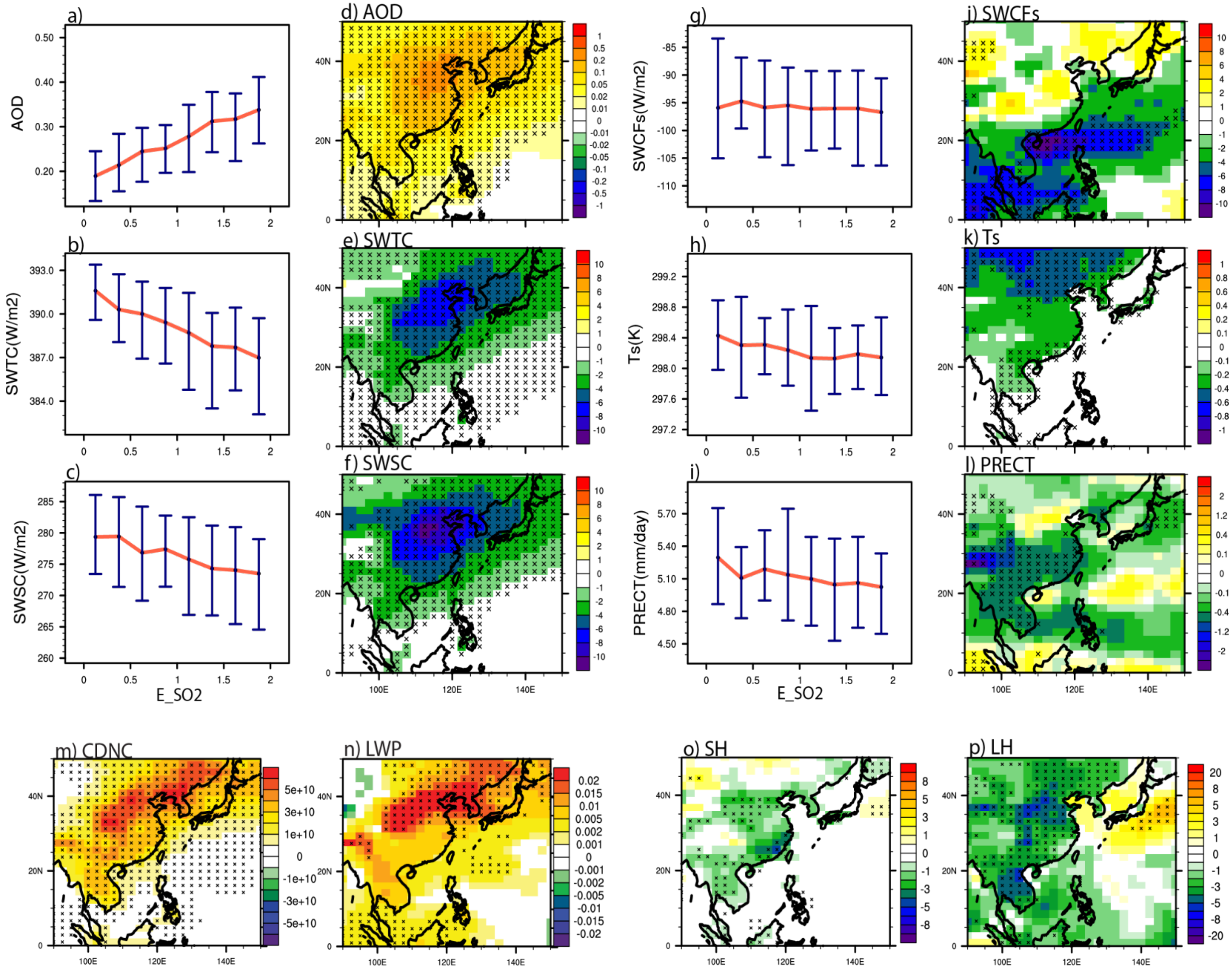
Pacific Northwest NATIONAL LABORATORY

Proudly Operated by **Battelle** *Since 1965*



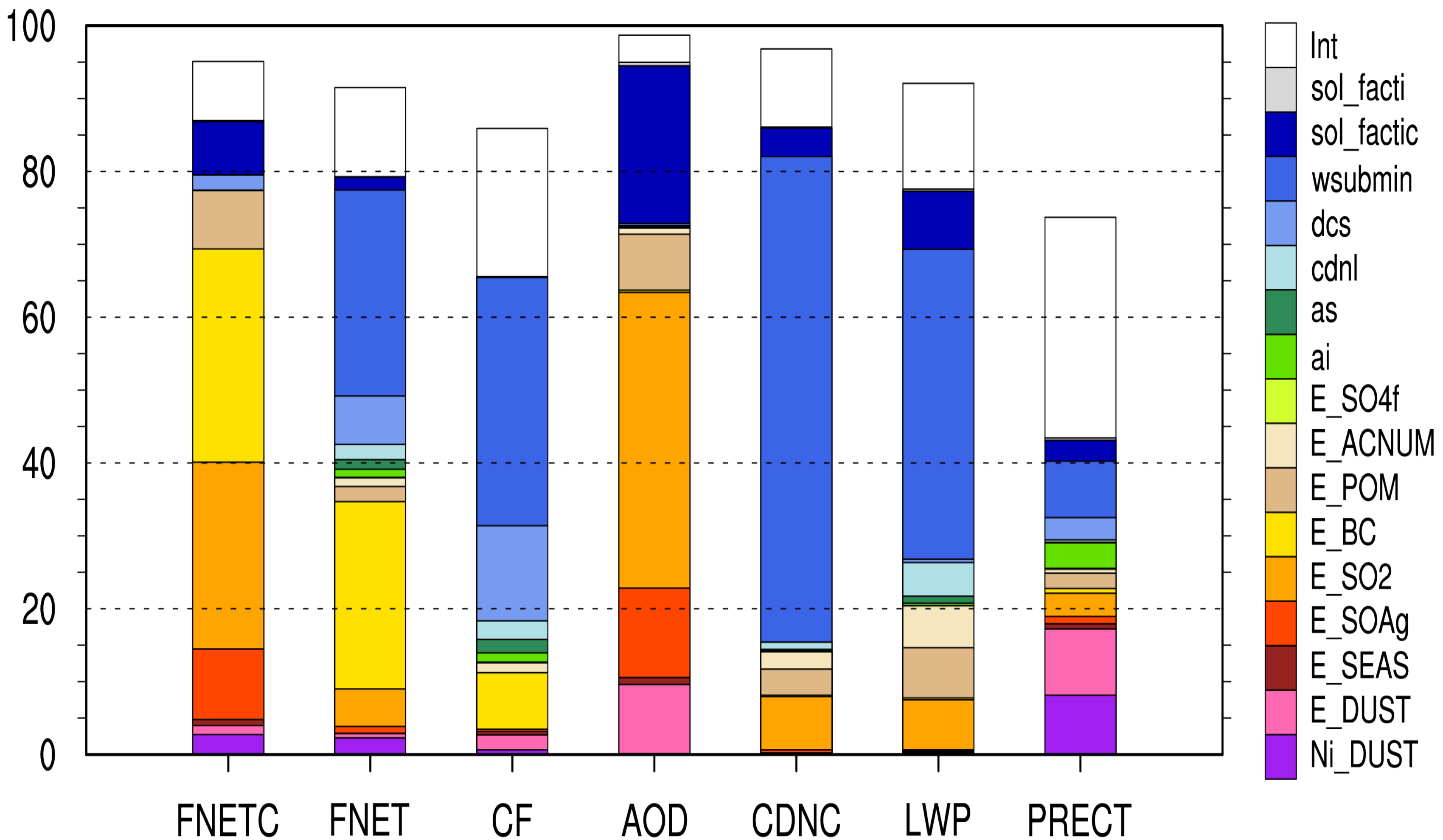
- PNNL: A-Ensemble
- Quasi Monte Carlo
- 16 parameters
- 256 sample sets (forward simulations)
- Each simulation: 5-yr
- Each parameters is sampled 256 times

Response to SO2 emission increase based on 256 simulations

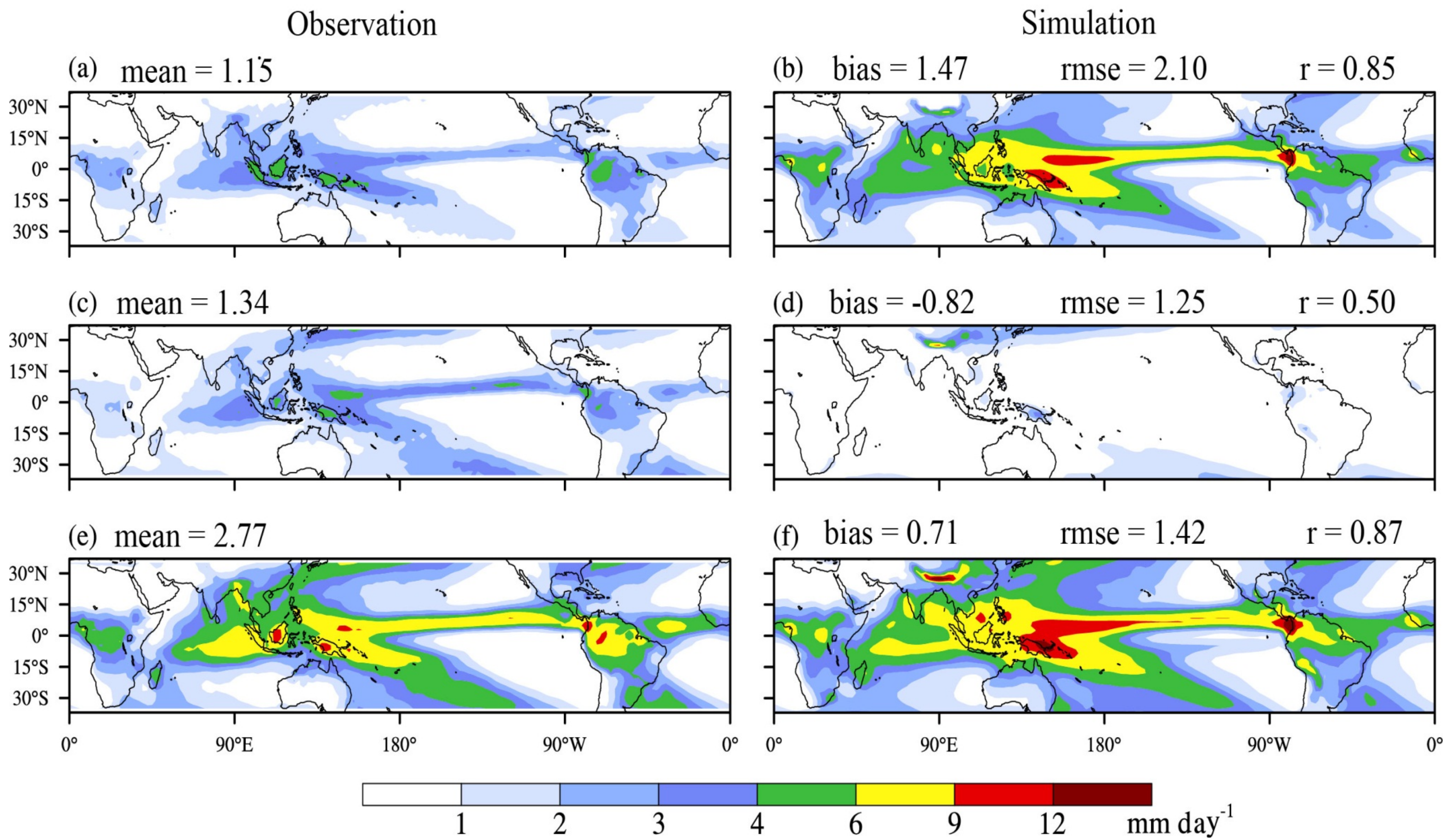


Fractional contributions of the 16 perturbed parameters to the total variance estimated by GLM

Pacific Northwest
NATIONAL LABORATORY
Proudly Operated by Battelle *Since 1965*



2. Convective Precipitation Calibration: CAM5 ZM scheme



Annual mean deep convective (top), stratiform (middle) and total (bottom) precipitation for 2001-2010 from TRMM/GPCP observation (left) and standard CAM5 (right)

Parameters in ZM scheme: Default, minimum and maximum values

Parameter	Default	Minimum	Maximum	Description[units]
C0_lnd	0.0059	0.001	0.045	Deep convective precipitation efficiency over land [m^{-1}]
C0_ocn	0.045	0.001	0.045	Deep convective precipitation efficiency over ocean [m^{-1}]
K_e	1.0E-06	0.5E-06	10E-06	Evaporation efficiency of precipitation [$(kg\ m^{-2}\ s^{-1})^{-1/2}\ s^{-1}$]
α	0.1	0.05	0.6	Maximum cloud downdraft mass flux fraction [fraction]
$CAPE_0$	70	20	200	Threshold value of CAPE for deep convection [$m^2\ s^{-2}$]
PE_lnd	-1.0E-03	-2.0E-3	0	Parcel fractional mass entrainment rate over land [m^{-1}]
PE_ocn	-1.0E-03	-2.0E-3	0	Parcel fractional mass entrainment rate over ocean [m^{-1}]
τ	3600	1800	28800	CAPE consumption time scale [s]
D_{ice}	25	10	50	Radius of detrained ice from deep convection [μm]

Evaluation Metric: Cost Function

$$E(\boldsymbol{m}) = \log\left[\frac{(\sigma_{\text{obs}}/\sigma_{\text{mod}} + \sigma_{\text{mod}}/\sigma_{\text{obs}})^2 (1+R_0)^k}{4(1+R)^k}\right]$$

MVFSA: Multiple Very Fast Simulated Annealing (a stochastic importance sampling algorithm)

$$m_i^{k+1} = m_i^{k} + y_i(m_i^{\max} - m_i^{\min})$$

$$y_i \in [-1,1]$$

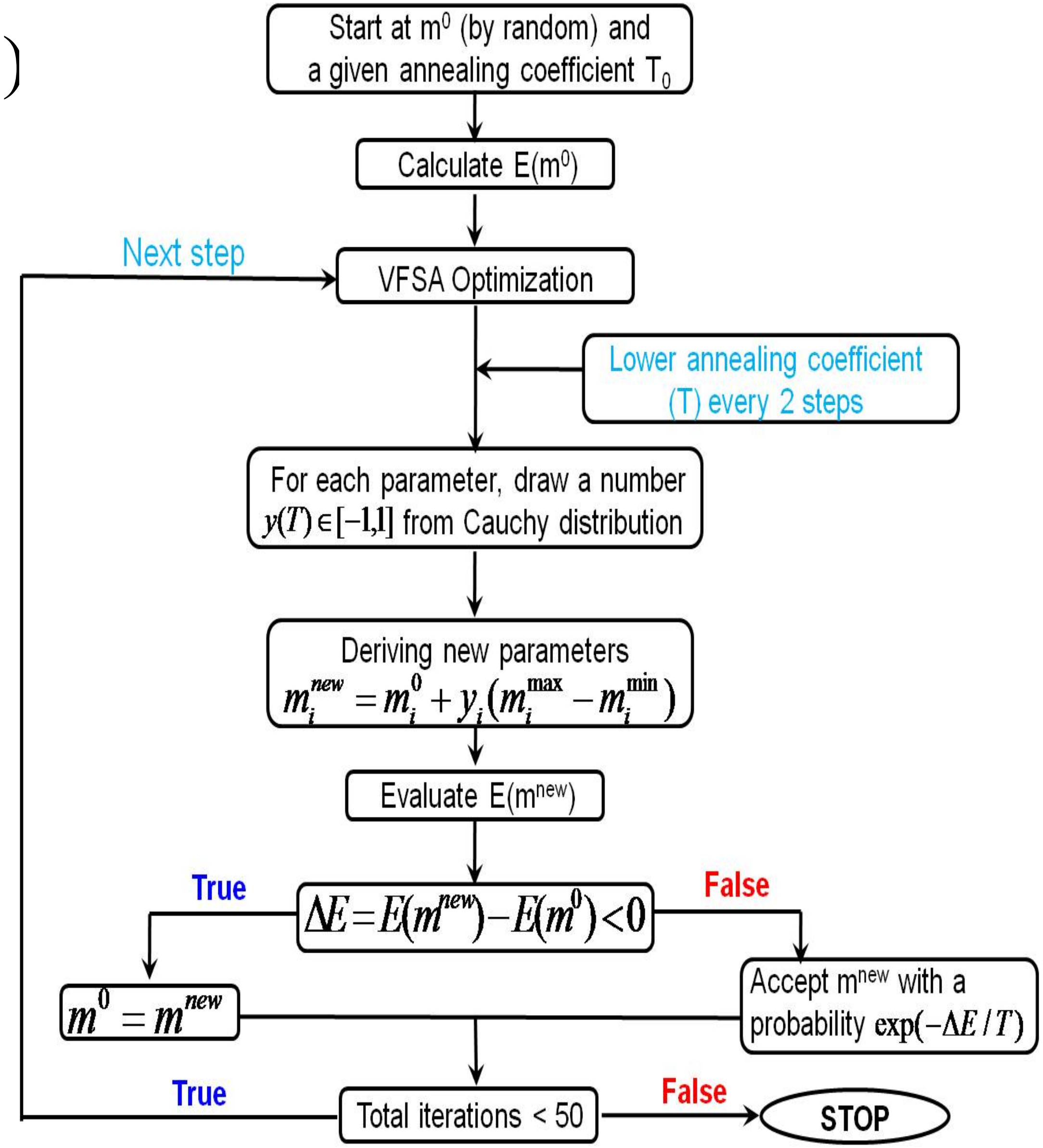
$$m_i^{\min} \le m_i^{k+1} \le m_i^{\max}$$

$$y_i = Sign(Random - 0.5)$$

$$\times T_k[(1 + \frac{1}{T_k})^{|2Random-1|} - 1]$$

$$T_k = T_0 \exp[-\alpha(k-1)^{1/NM}]$$

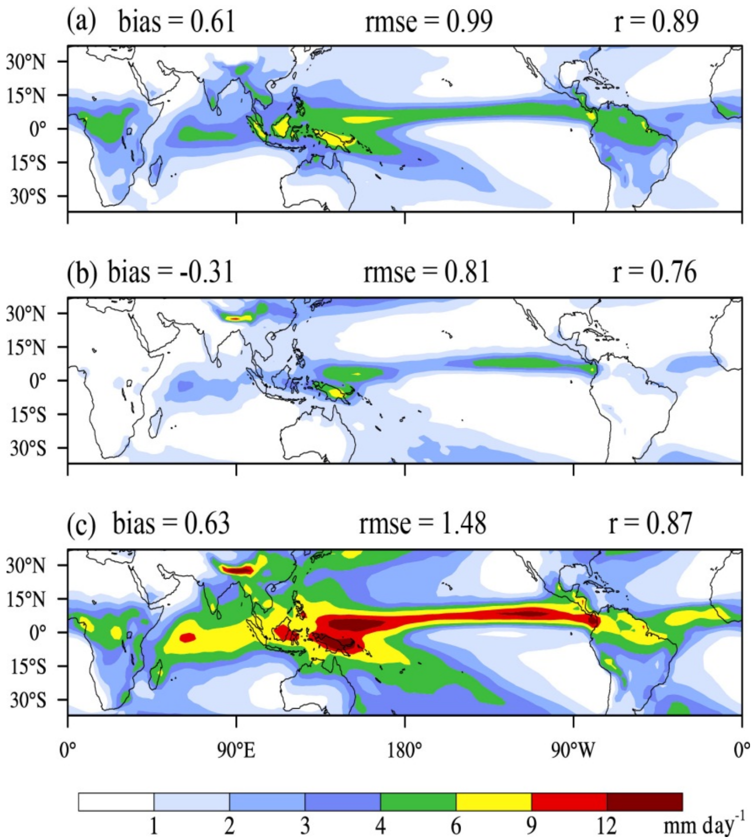
(Jackson et al, J. Climate, 2004)

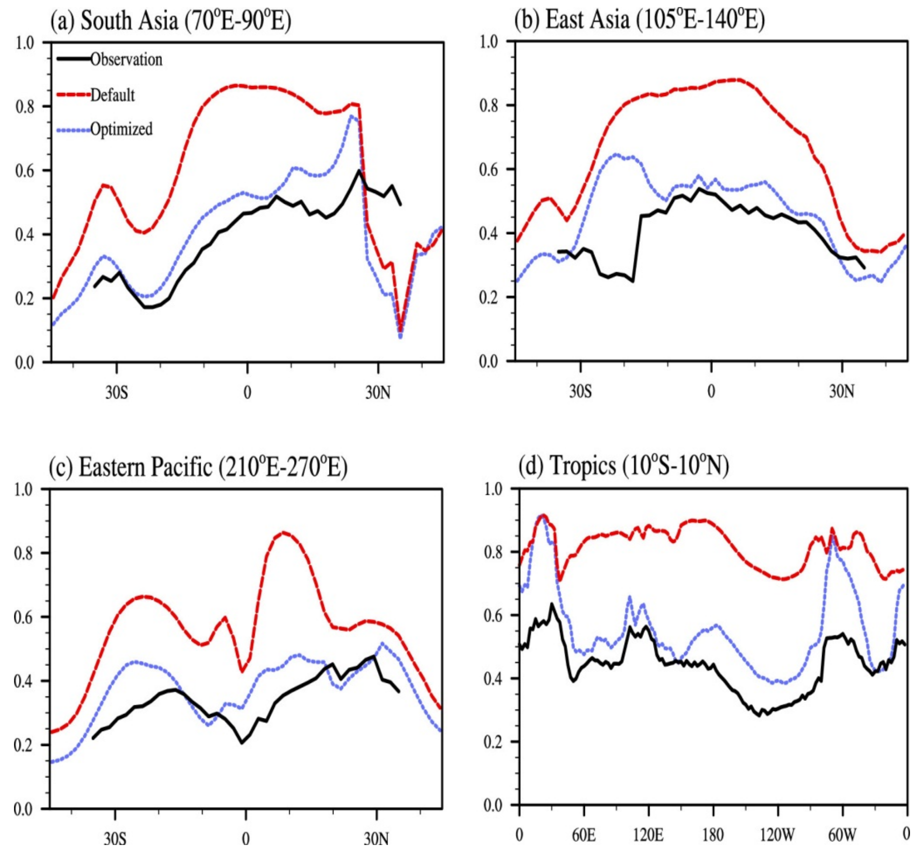


Optimized Results

Annual mean deep convective (top), stratiform (middle) and total (bottom) precipitation simulated by CAM5 with the optimized parameters.

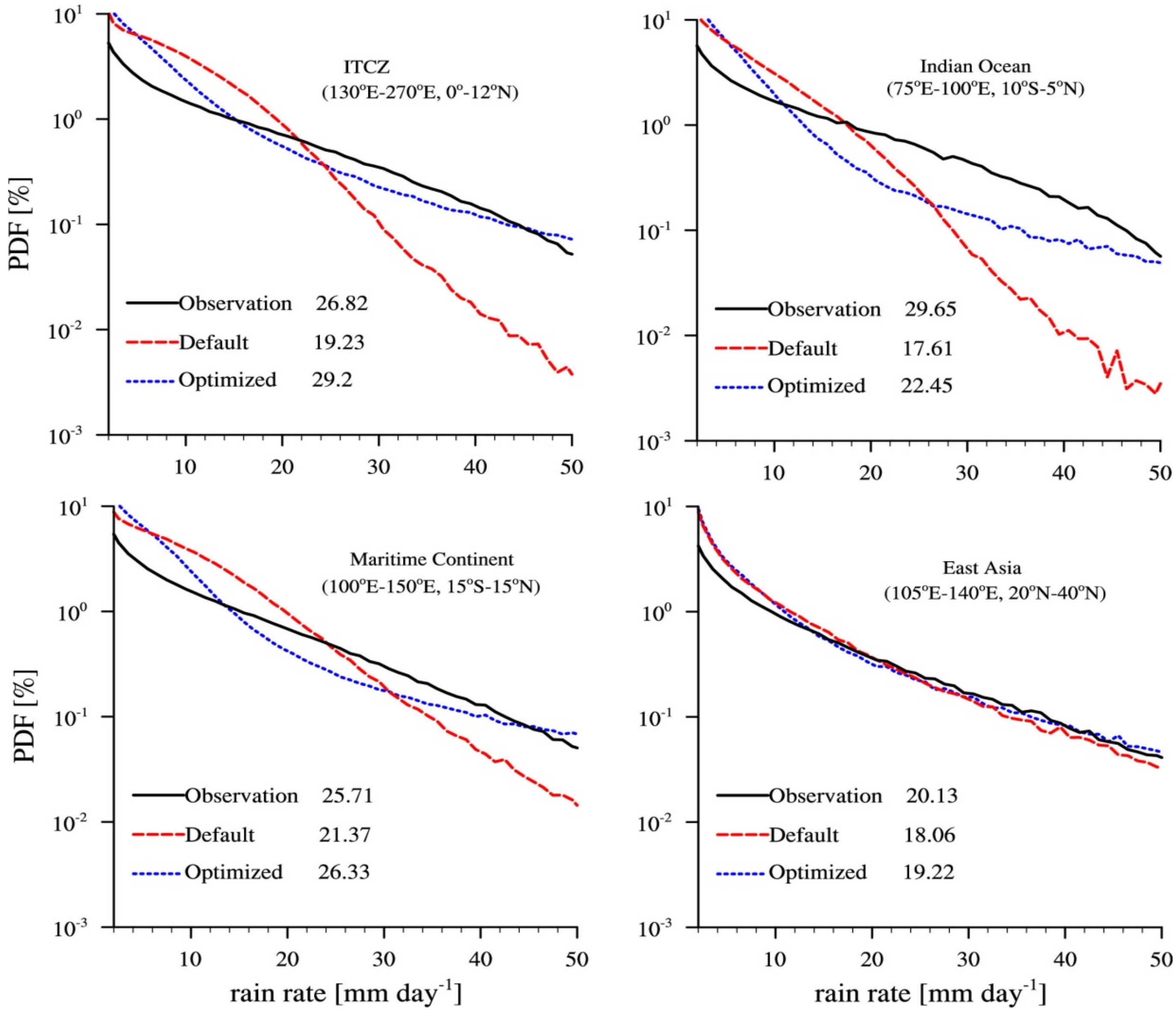
Meridional distributions of the ratios of deep convective vs. total precipitation from the observation (black) and two model simulations over four regions.





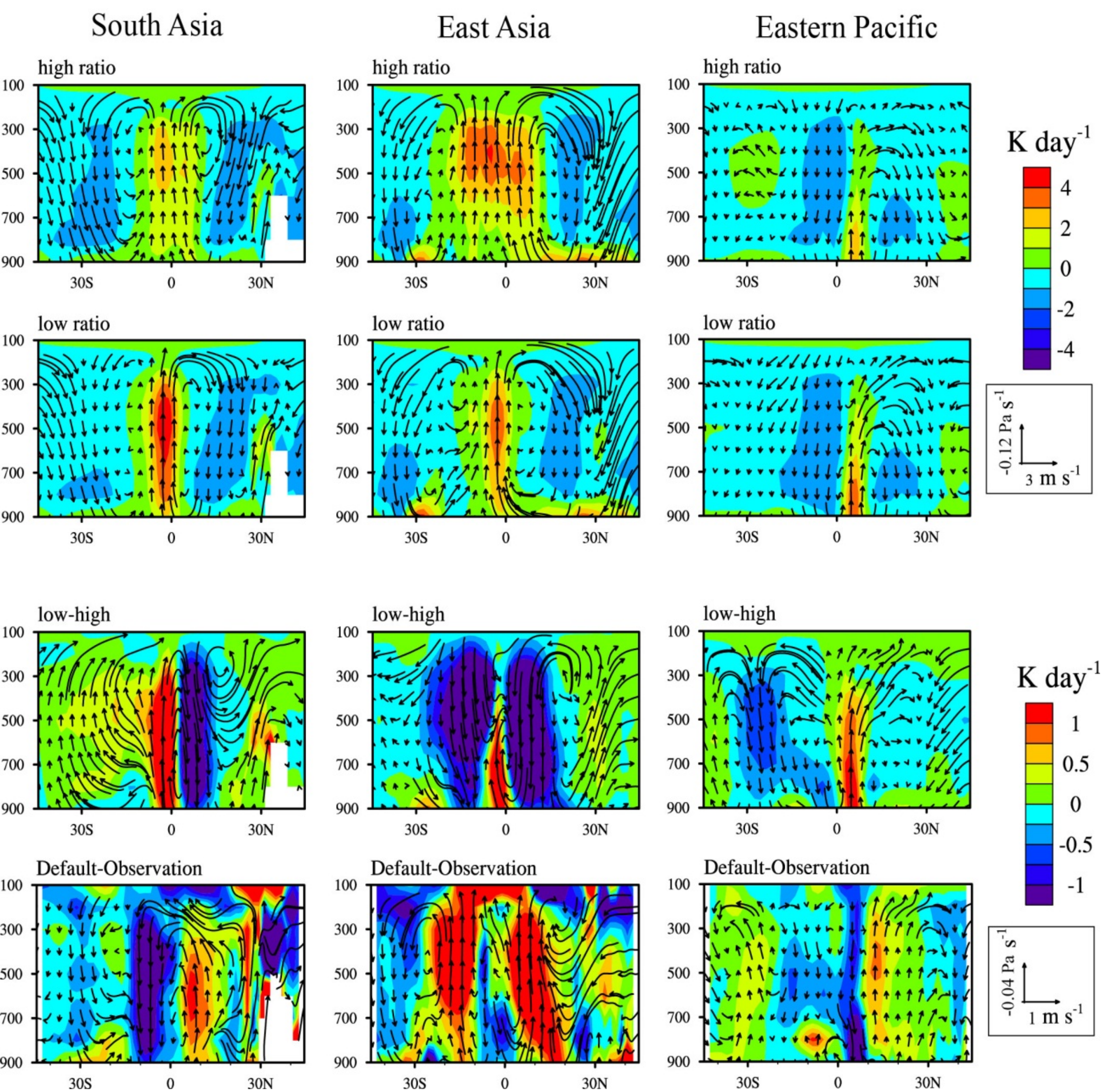
Frequency distributions of daily precipitation as a function of rain rates. The rain rates at 95th percentile are also given

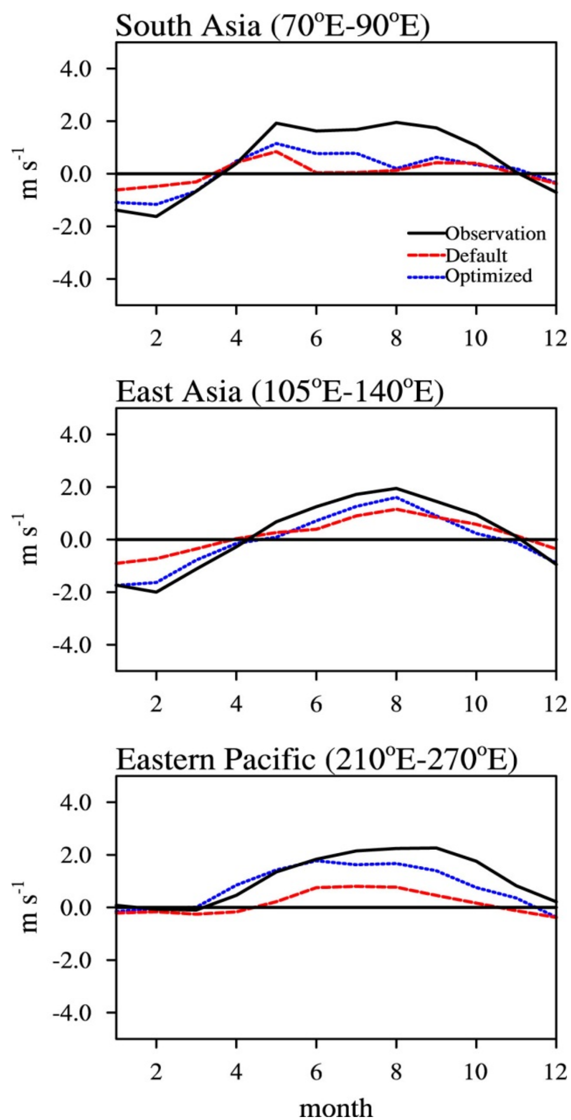
Pacific Northwest
NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965



Impact on Circulation

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965





Calibrating the Convective Precipitation for the Community Atmosphere Model (CAM5)

Pacific Northwest NATIONAL LABORATORY

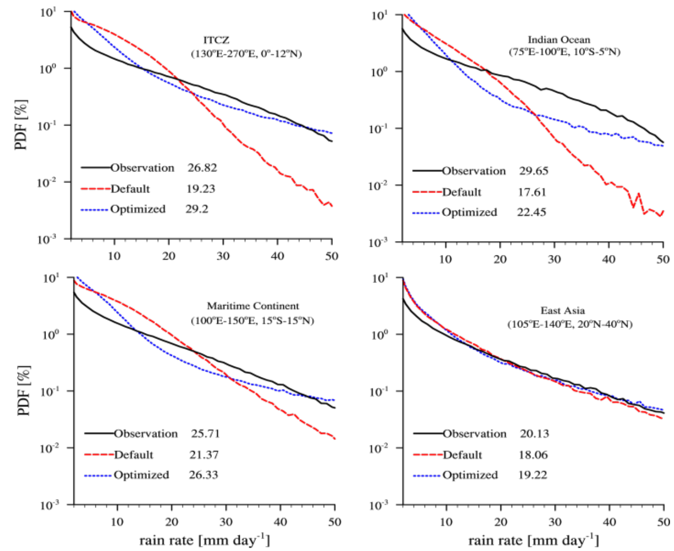
Proudly Operated by Battelle *Since 1965*

Objective

- Calibrate the convective precipitation in the global climate model CAM5 and study the subsequent impact of improved convection.

Approach

- Applied an Uncertainty Quantification (UQ) technique to improve Zhang-McFarlane (ZM) deep convection scheme in CAM5.
- Examined the sensitivity of precipitation and circulation to key parameters in the ZM scheme in CAM5, using a stochastic importance-sampling algorithm.
- Evaluated the subsequent impact of improved deep convection on the global circulation and climate.



Frequency distributions of daily precipitation as a function of rain rates observed and simulated with the standard and optimized parameters over four regions. The numbers denote the rain rates at the 95th percentile.

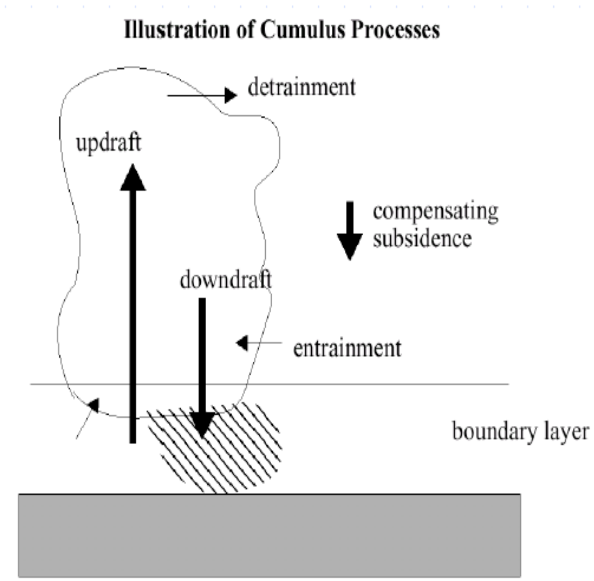
Impact

- Simulated convective precipitation is most sensitive to the parameters of Convective Available Potential Energy (CAPE) consumption timescale, parcel fractional mass entrainment rate, and maximum downdraft mass flux fraction.
- As the optimal parameters are used, positive impacts on some aspects of the atmospheric circulation and climate are found, including mitigated double ITCZ problem, improved East Asian monsoon precipitation, and annual cycles of the cross-equatorial jets.

Yang B, Y Qian, G Lin, LR Leung, PJ Rasch, GJ Zhang, SA McFarlane, C Zhao, Y Zhang, H Wang, M Wang, and X Liu . 2013. “Uncertainty quantification and parameter tuning in the CAM5 Zhang-McFarlane convection scheme and impact of improved convection on the global circulation and climate.” Journal of Geophysical Research 118. DOI:10.1029/2012JD018213.

A regional model study on K-F convection scheme in WRF model

Pacific Northwest
NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965



1. Pd: Downdraft Mass Flux Rate
2. Pe: Environmental Air Entrainment Rate
3. Pt: maximum TKE in sub-cloud layer
4. Ph: starting height of downdraft above cloud base
5. Pc: average consumption time of CAPE

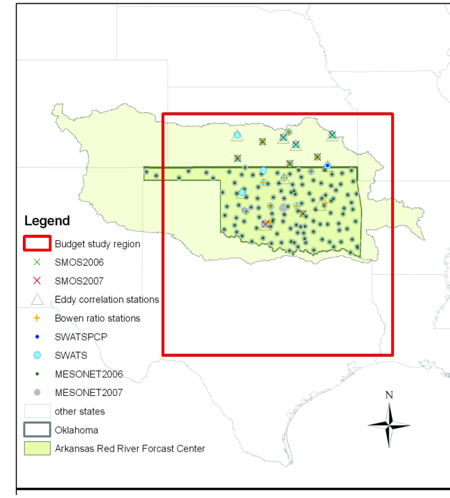
$$DMF/UMF = 2\times(1-RH)\boxed{\times 2^{Pd}, Pd\in(-1,1)}$$

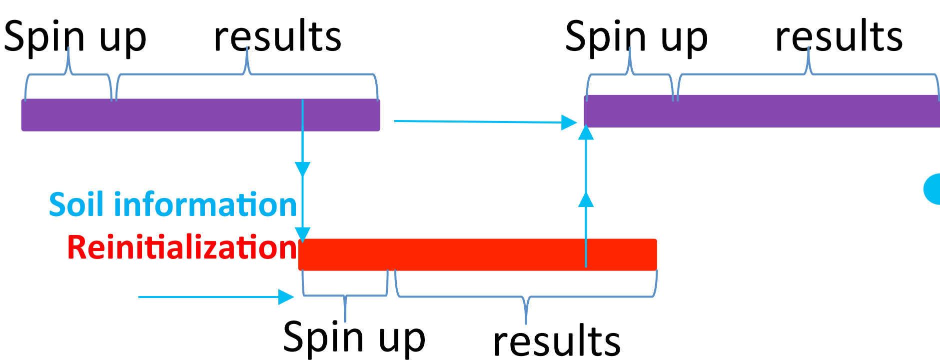
$$\delta M_e / M_{u0} = \frac{-0.03\cdot\delta p}{R}\boxed{\times 2^{Pe}, Pe\in(-1,1)}$$

Parameter	Default	Minimum	Maximum	Description
Pd	0	-1	1	Coefficient related to the downdraft mass flux rate
Pe	0	-1	1	Coefficient related to Entrainment mass flux rate
Pt	5	3	12	maximum TKE in sub-cloud layer (m^2 s^{-2})
Ph	150	50	350	starting height of downdraft above cloud base (hPa)
Pc	2700	900	7200	averaged consumption rate of CAPE (s)

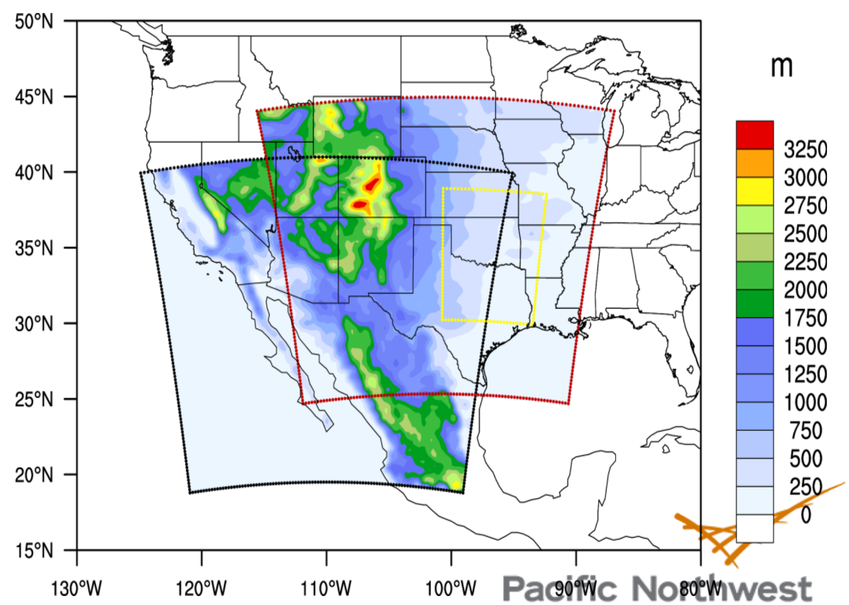
Model configuration and experiment design

- South Great Plain (25°N-44°N, 112°W-90°W)
- Resolution: 25 KM
- Simulation period: May 1 to July 1, 2007
- Analysis: June 2007
- UW 1/8-degree gridded daily data (P, T, Wind)





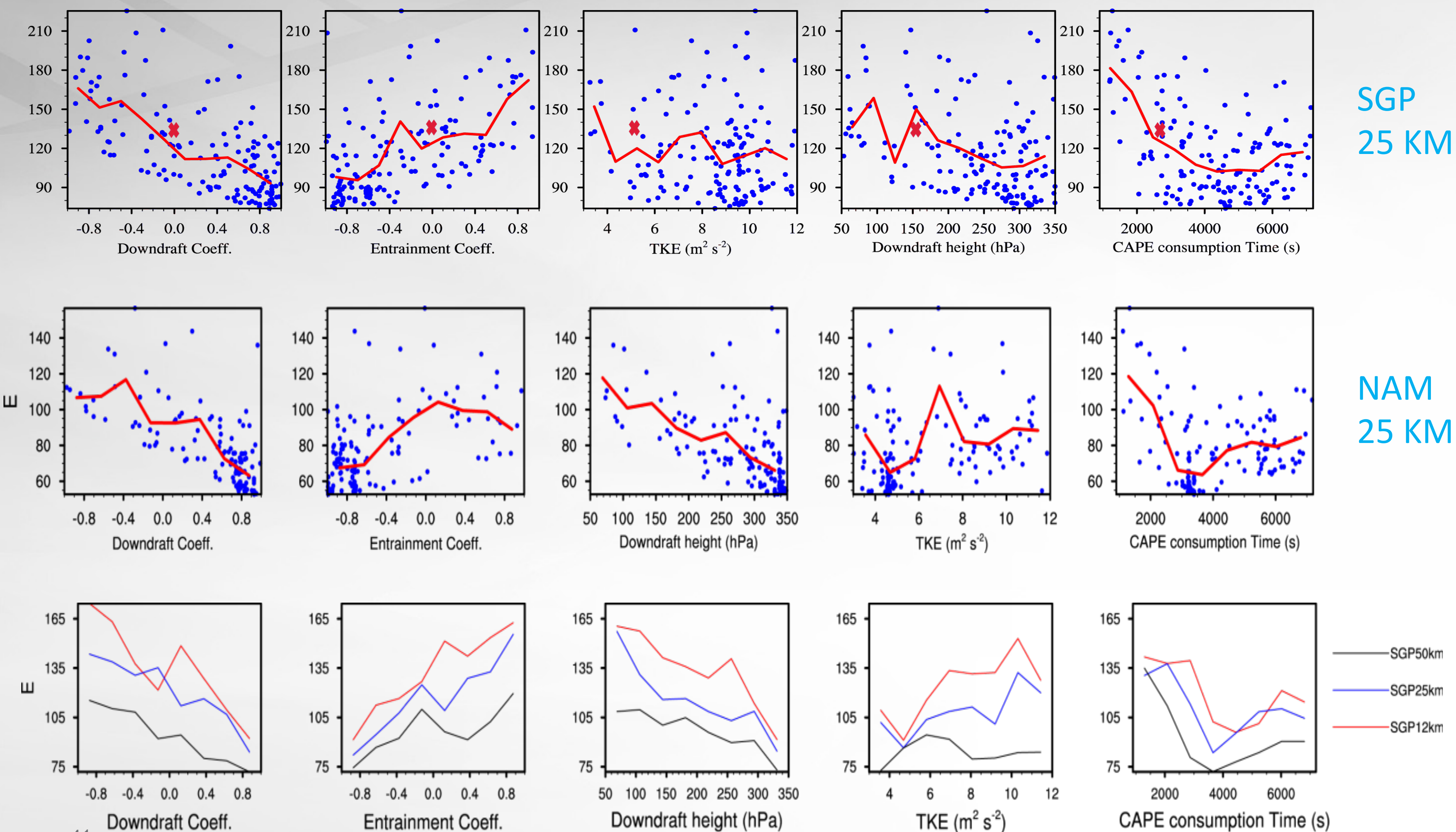
Each simulation:3-day
1-day spin up time
2-day results are used



- Microphysics: Morrison vs. WSM6
- Radiation: RRTMG vs. CAM
- Planetary boundary layer: MYJ
- Surface physics: Noah Scheme
- Cumulus parameterization: **New K-F cumulus scheme**

Pacific Northwest
NATIONAL LABORATORY

Transferability of sensitivity and parameter tuning across physical processes, spatial scales, and climatic regimes (Yan et a., 2013)



Uncertainty quantification and parameter tuning: a case study of convective parameterization in WRF

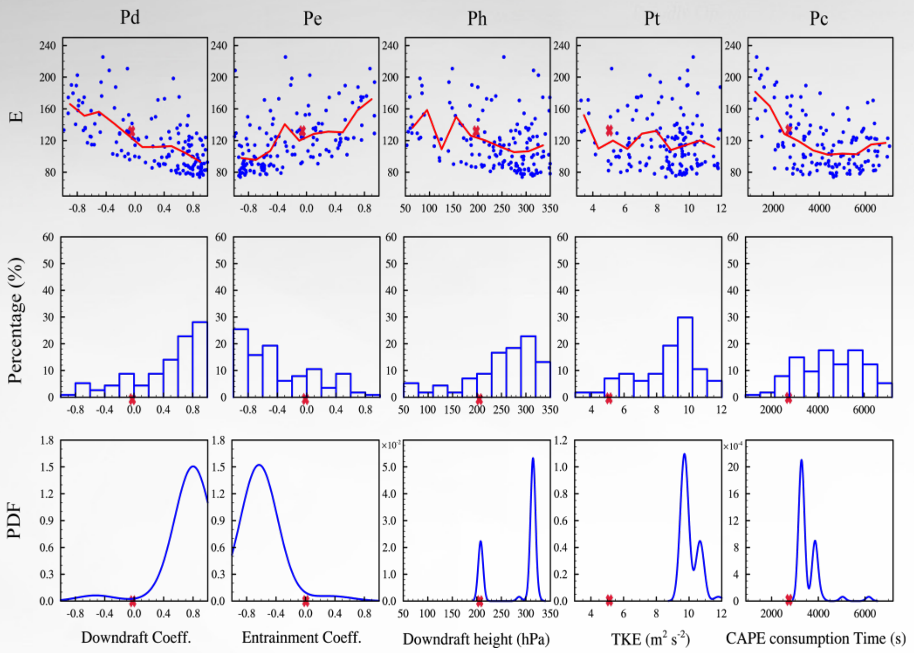
Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle *Since 1965*

Objective

- Use observational data to calibrate input parameters in convection scheme and validate results of WRF
- Explore UQ and parameter tuning across physical processes, spatial scales and climatic regimes

Research

- Focus on five key input parameters in Kain-Fritsch used in WRF
- Use MVFSA, a stochastic importance sampling algorithm, to minimize model errors
- Apply optimized parameters for precipitation simulation to another spatial resolution and to another region with a different climate regime.



(Top) The response of model performance (quantified as E) to five input parameters. (Middle) The frequency distributions of “good” experiments as a function of each parameter. (Bottom) The marginal probability density functions (PDF) for the five input parameters derived by kernel density estimation.

Impact

- Precipitation bias in model significantly reduced when using five optimal parameters identified by MVFSA
- Identified benefits of optimal parameters that are transferable across processes, scales and climatic regimes.

Yang B, Qian Y, Lin G, Leung R, and Zhang Y. 2012. “Some Issues in Uncertainty Quantification and Parameter Tuning: A Case Study of Convective Parameterization Scheme in the WRF Regional Climate Model,” Atmospheric Chemistry and Physics, 12, 2409-2427, doi: 10.5194/acp-12-2409-2012.

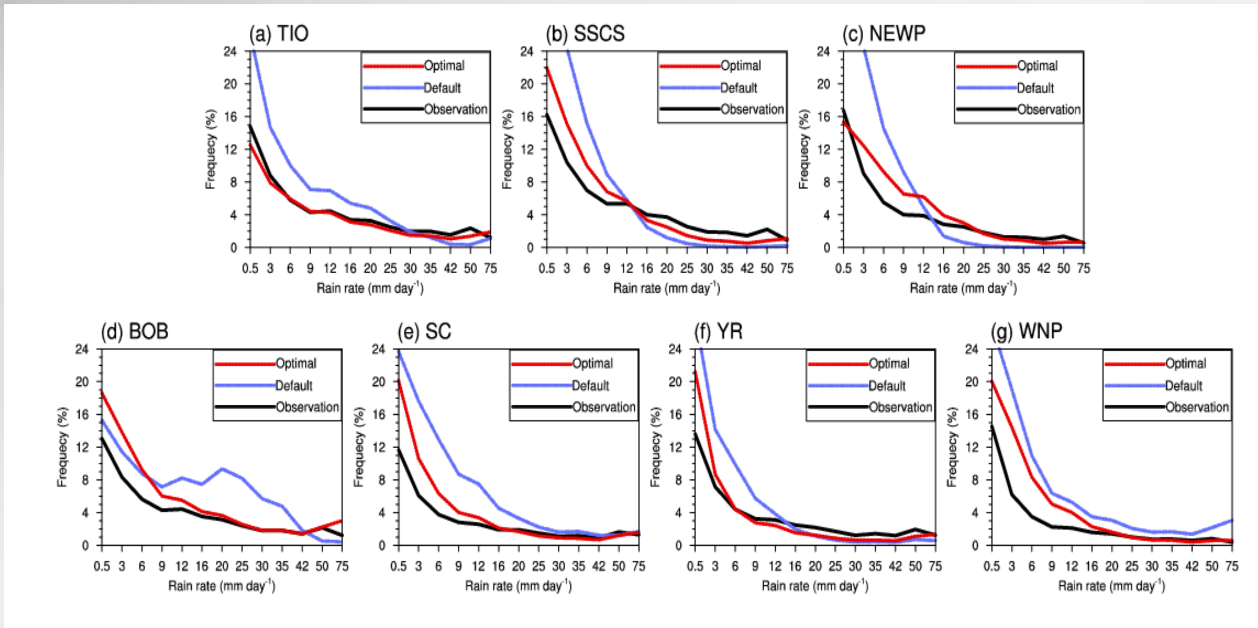
Parameter Tuning and Calibration of RegCM3 with MIT–Emanuel Cumulus Parameterization Scheme over CORDEX East Asia Domain (Zou et al., 2014)

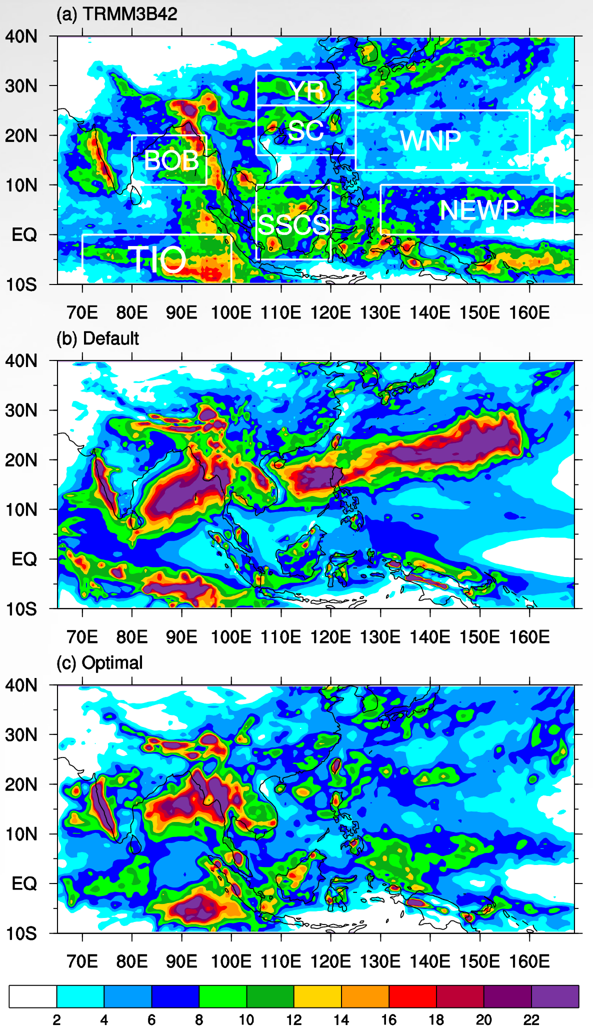
Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

Parameter	Default	Minimum	Maximum	Description
RHC	—	0.4	0.9	Convection is activated when the RH averaged from the cloud top to the cloud base is larger than a critical value (RHC). In the default setting, the convection is driven by the buoyancy, and effects of the large-scale environment are not considered.
$C_{asc\text{–}land}$	0.4	0.2	0.8	Autoconversion scale factor over ocean
$C_{asc\text{–}ocean}$	0.4	0.2	0.8	Autoconversion scale factor over ocean
RH_{min_land}	0.8	0.6	1.0	Gridbox RH threshold for cloudiness over land
RH_{min_ocean}	0.9	0.6	1.0	Gridbox RH threshold for cloudiness over ocean
α	0.2	0.0002	0.8	Rate at which the cloud base upward mass flux is relaxed to steady state
l_0	0.0011	0.0001	0.05	Amount of cloud water available for precipitation conversion

Parameter Tuning and Calibration of RegCM3 with MIT–Emanuel Cumulus Parameterization Scheme over CORDEX East Asia Domain (Zou et al., 2014)

Pacific Northwest
NATIONAL LABORATORY
Proudly Operated by Battelle *Since 1965*





Summary

1. Uncertainty Quantification (UQ) methods, such as forward and inverse modeling, can be useful tools for

- better understanding the model behavior and physical processes to help model development
- guiding hand-on model tuning and calibrating
- quantifying the impact of anthropogenic forcing (e.g. aerosol) in a more robust way

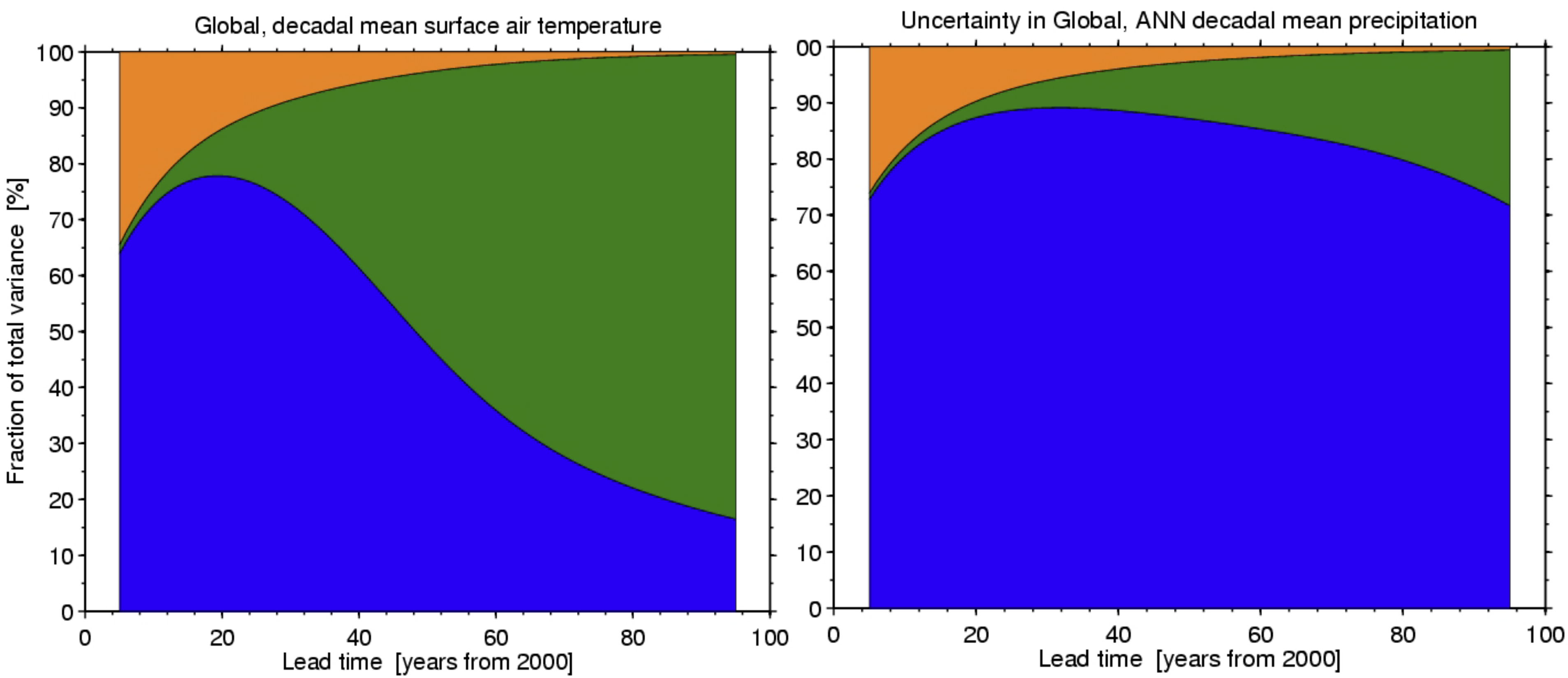
2. Process-level calibration (e.g. cloud, convection) is critical for getting right response (e.g. P, T) for right physical reason.

3. Transferability of sensitivity and parameter tuning across physical processes, spatial scales, and climatic regimes need to be further investigated.

4. More efficient strategies for model evaluation and tuning need to be further explored (e.g. short simulations, nudging simulations, surrogate model)

Sources of uncertainties

(Hawkins and SuHon 2009, BAMS)



Yellow = internal variability

Green = scenario uncertainty

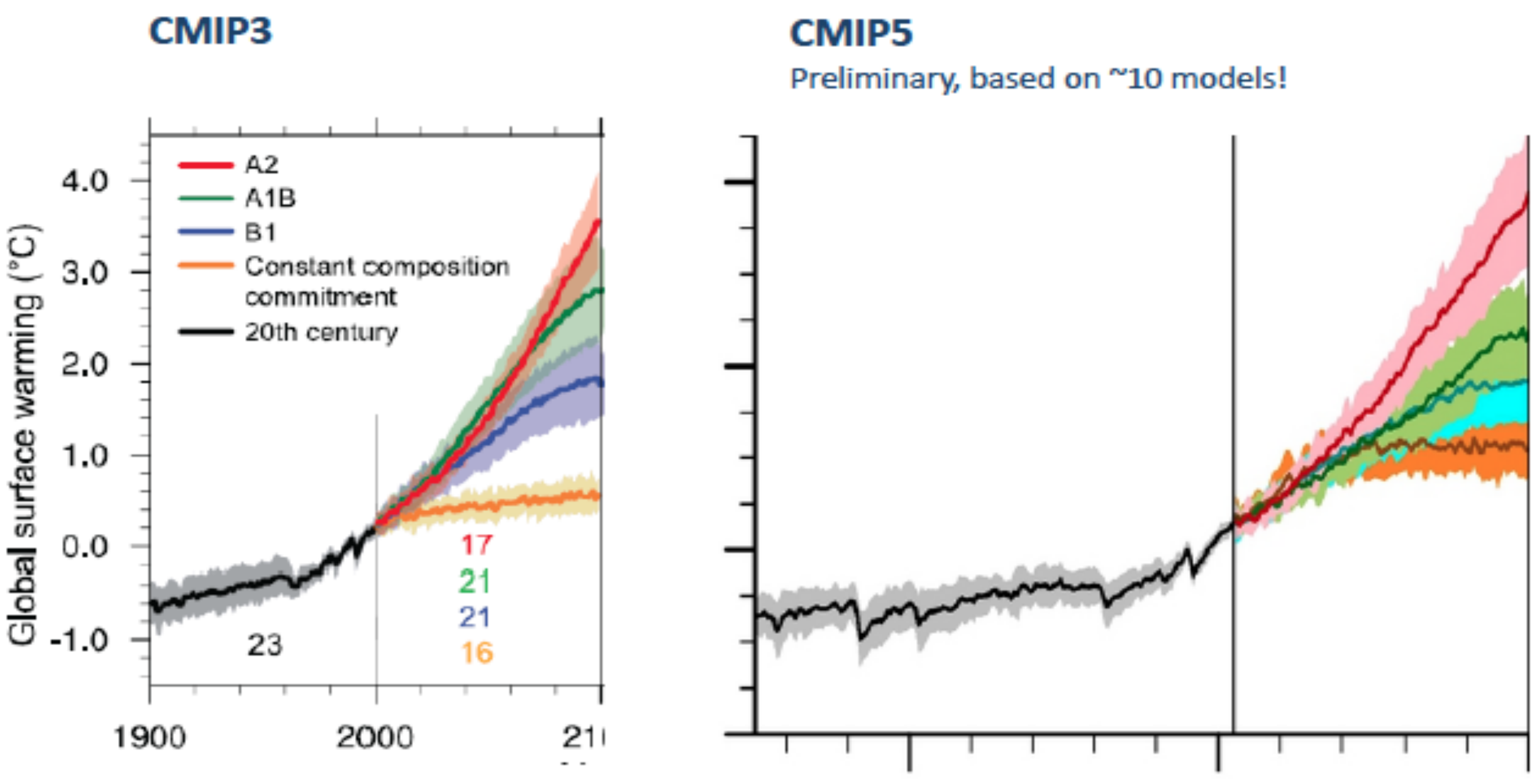
Blue = model uncertainty

Δ = mod params + mod struct. error + obs uncert.

Reducing the uncertainty of climate modeling and projection is difficult, if not impossible

Pacific Northwest
NATIONAL LABORATORY
Proudly Operated by Battelle *Since 1965*

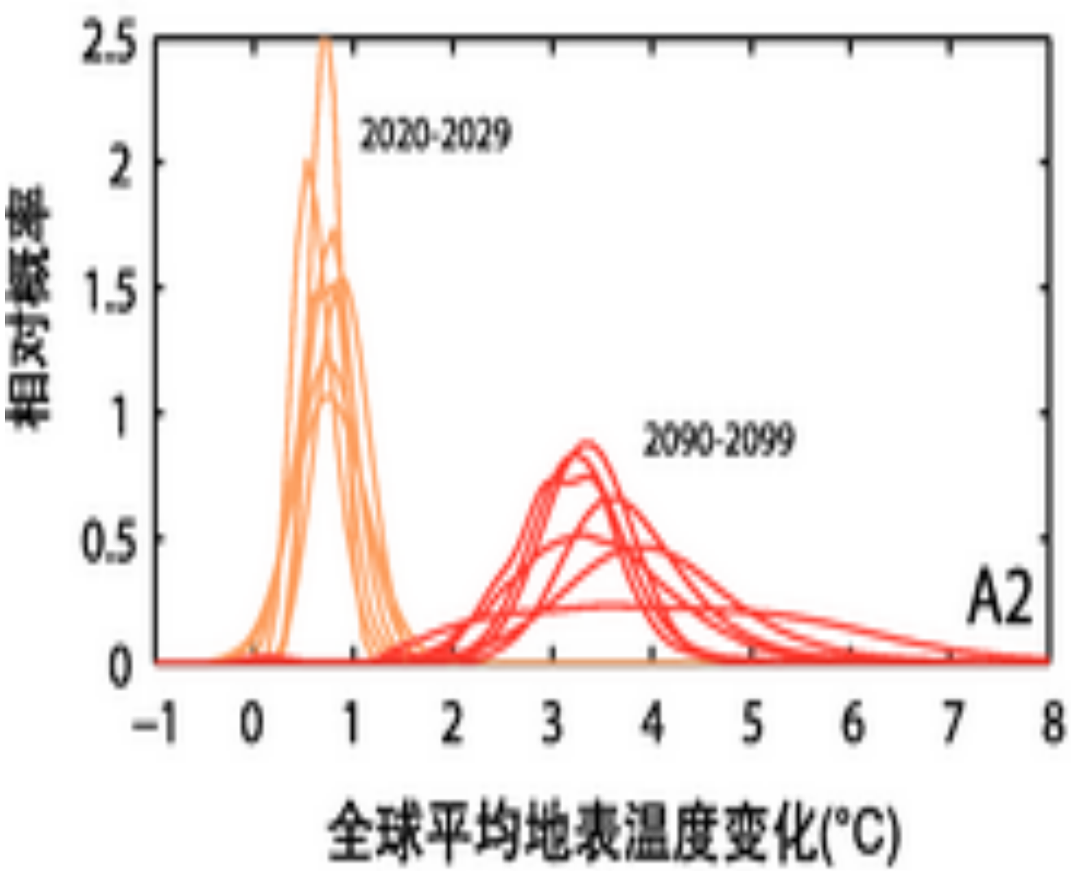
Model agreement with observations improves, but future spread is not decreasing.



Relating present day climate to future changes (Telbaldi et al., 2012)

What quantifying the uncertainty is important ?

Pacific Northwest
NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965



IPCC AR4, 2007

- ❑ Reality in climate modeling
- ❑ Difficulties in reducing uncertainty
- ❑ Need for impact and adaptation community
- ❑ Strategic and political rationale

Applications of UQ in climate modeling

1. Sensitivity Analysis (SA)
Response of climate to uncertain input parameters

2. Surrogate models
Climate model emulators for UQ and optimization
GLM, GPM (Gaussian Process Models), RF (Random Forests)

3. Model calibration
Tuning of uncertain parameters using observations
(validation, verification, optimization, observation data uncertainty)

4. Forward UQ
Characterize predictive accuracy from UQ
(constructing ensembles, uncertainty propagation, present-future extrapolation)

Sensitivity of phase of PDC to C-Ensemble parameters

Pacific Northwest NATIONAL LABORATORY

Proudly Operated by Battelle *Since 1965*