BERKELEY LAB

Lawrence Berkeley National Laboratory

U.S. DEPARTMENT OF ENERGY

UNIVERSITY OF CALIFORNIA

BERKELEY LAB

LAWRENCE BERKELEY NATIONAL LABORATORY

U.S. DEPARTMENT OF ENERGY

# Parameter Uncertainty for Attribution Studies and Climate Extremes


**Ben Timmermans[1,2] (ben.timmermans@gmail.com)**

**William Collins[2], Travis O'Brien[2], Dáithí Stone[3,2] & Mark Risser[2]**

*[1] National Oceanography Centre (Southampton, UK)*

*[2] Lawrence Berkeley National Laboratory (Berkeley, US)*

*[3] National Institute of Water and Atmospheric Research (New Zealand)*



EGU General Assembly 2020 (Virtual)

7th May 2020

# Outline

- Motivation for a PPE for extremes and attribution studies
- Experimental description
  - Numerical model
  - Experimental design
  - Emulation / Sensitivity analysis details
- Sensitivity analysis results
- Conclusions (so far)

# Motivation

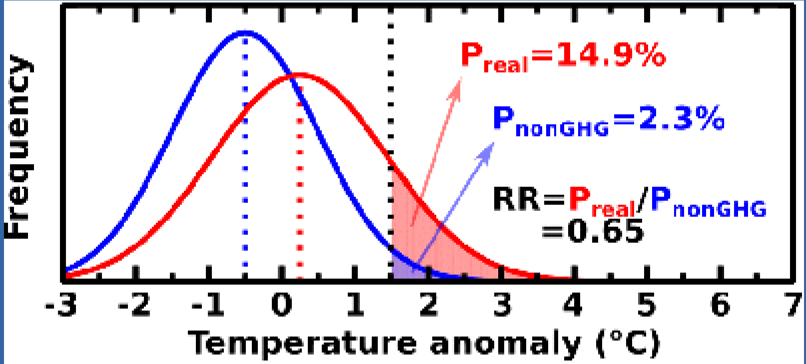
- Extreme event attribution studies, such as the CLIVAR C20C+ project **(Stone et al., 2019)**, often rely heavily on numerical models (e.g. Community Earth System Model, CESM).
- Modeling uncertainty remains poorly understood in the context of simulated extremes and attribution studies.
- Perturbed physics within the CESM is know to affect weather and climate extremes **(e.g. Qian et al. 2015, Kooperman et al., 2018)**
- Important questions:

  i) Are extremes (precip. / temp. / other) sensitive to perturbed physics, and would sensitivity lead to increased uncertainty in attribution statements, such as those based on risk ratios?

  ii) Can we explore this issue with a large ensemble, and apply surrogate modeling / machine learning methods to formally this question?

# Output extremes for event attribution

- Extreme event attribution studies often compute a "risk ratio" **(e.g. Angelil et al., 2017)**.




- Clearly, uncertainty induced in either the **red** or **blue** histograms could affect the mean estimate of the risk ratio (RR), or its uncertainty.
- Estimation of the tails of these histograms can be done in a number of ways: counting events, high quantiles of continuous variables, fitted parameterised extreme value distributions.

# Parameterisation uncertainty in extremal output

- We proposed to design and run a large perturbed physics ensemble using CESM, largely following **Qian et al. 2015**, in order to evaluate the effects of uncertain physics parameterisations on estimation of relevant extremes, and associated risk ratios.
- The ensemble predominantly spans the period of time (2010 to 2013) for a number of events (~50) described in the early attribution literature, typically reported in BAMS **(see e.g. Angelil et al., 2017)**
- The ensemble was run during 2017 / 2018 at the U.S. Department of Energy's **National Energy Research Scientific Computing Center (NERSC https://www.nersc.gov/)**

# PPE: Design

# Design Summary

- CAM5 / CLM4 (AMIP) @ ~1 degree global spatial resolution
- C20C+ configuration ("All-Hist" + "Nat-Hist" runs), **Stone et al. (2019)**
- Physics perturbations via atmospheric convection processes (deep / shallow convection, cloud fraction). Parameters on next slide.
- 150 PPE members, spanning Oct 2010 to Oct 2013 and comprising ~28 to 42 initial condition (IC) sub-member realisations (Oct 2010 to Oct 2013).
- Each PPE member therefore comprises between ~50 and 100 years and approximately 11,000 years per scenario (All-Hist + Nat-Hist runs).

2 years spin-up
Nat-Hist (1) / All-Hist (1)
Default parameters

2007 -> 2009
Spin-up Nat-Hist
(1 member)

2009 / 01 -> 2010 / 10
(150 PPE members)

2010 / 10 -> 2013 / 10
(150 PPE x 28 IC)

2007 -> 2009
Spin-up All-Hist
(1 member)

2009 / 01 -> 2010 / 10
(150 PPE members)

2010 / 10 -> 2013 / 10
(150 PPE x 28 IC)

**= 2 x 11,000 years**

# Experiment design: CAM parameters

| Parameter name | Description | Default value | Physics scheme |
|---|---|---|---|
| tau | Time scale for consumption rate of deep CAPE. | 3600 s | ZM deep convection |
| dmpdz | Parcel fractional mass entrainment rate. | -1.0e-3 | ZM deep convection |
| c0_lnd | Deep convection precipitation efficiency over land. | 5.9e-3 | ZM deep convection |
| ke | Evaporation efficiency of precipitation. | 1.0e-6 | ZM deep convection |
| rhminl | Minimum relative humidity for low stable clouds. | 0.80 | Cloud fraction |
| criqc | Maximum updraft condensate. | 0.70 | UW shallow convection |
| kevp | Evaporation efficiency. | 2.0e-6 | UW shallow convection |

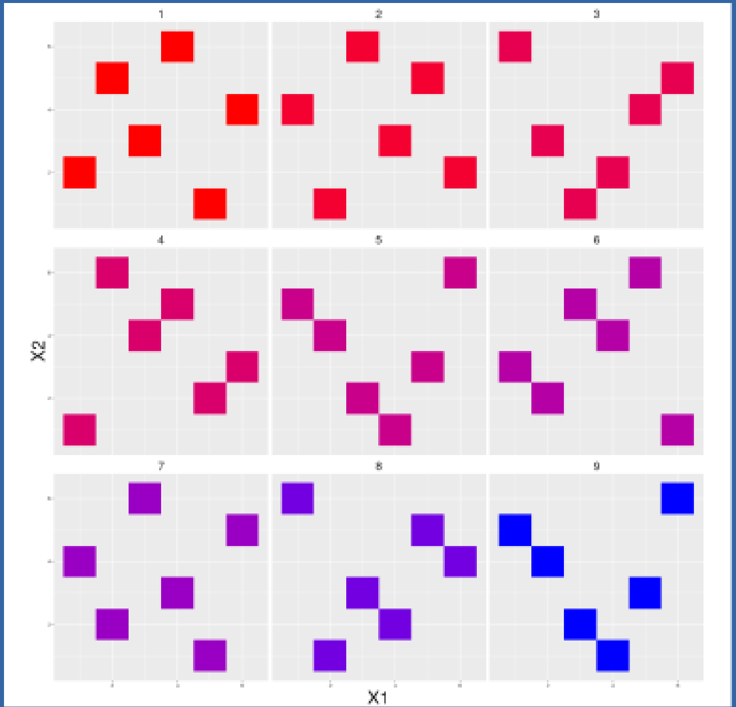
- Parameters known to induce variability (E.g. Qian et al. 2015)

# Experiment design: Hypercube configuration

- Latin hypercube (7-D) over parameters.
- Allows for "design expansion" if new points in the design space are required later (e.g. for greater space filling or validation).
- Figure to the left: an example of 2-D "k-extended Latin hypercube" **(Williamson, 2015)**

  2-variable design comprises 9 mutually orthogonal hypercubes.

  Not all sub-hypercubes are necessarily required thus a "staged" approach can optimise resources.



# Examples of output visualisation:
# Total precipitation (PRECT)

# Visualisation summary

- Projection of global output (DJF, 2011 / 2012) against parameters reveals structural dependence (or lack of dependence) on parameterisation.
- The following slides (1-4) show projections of output (PRECT, y-axis: mean, variance, Q95, max) against parameter “tau” (CAPE timescale, x-axis) for:

  **1) Global ocean surface**

  **2) Global land surface**

  **3) South American continent**

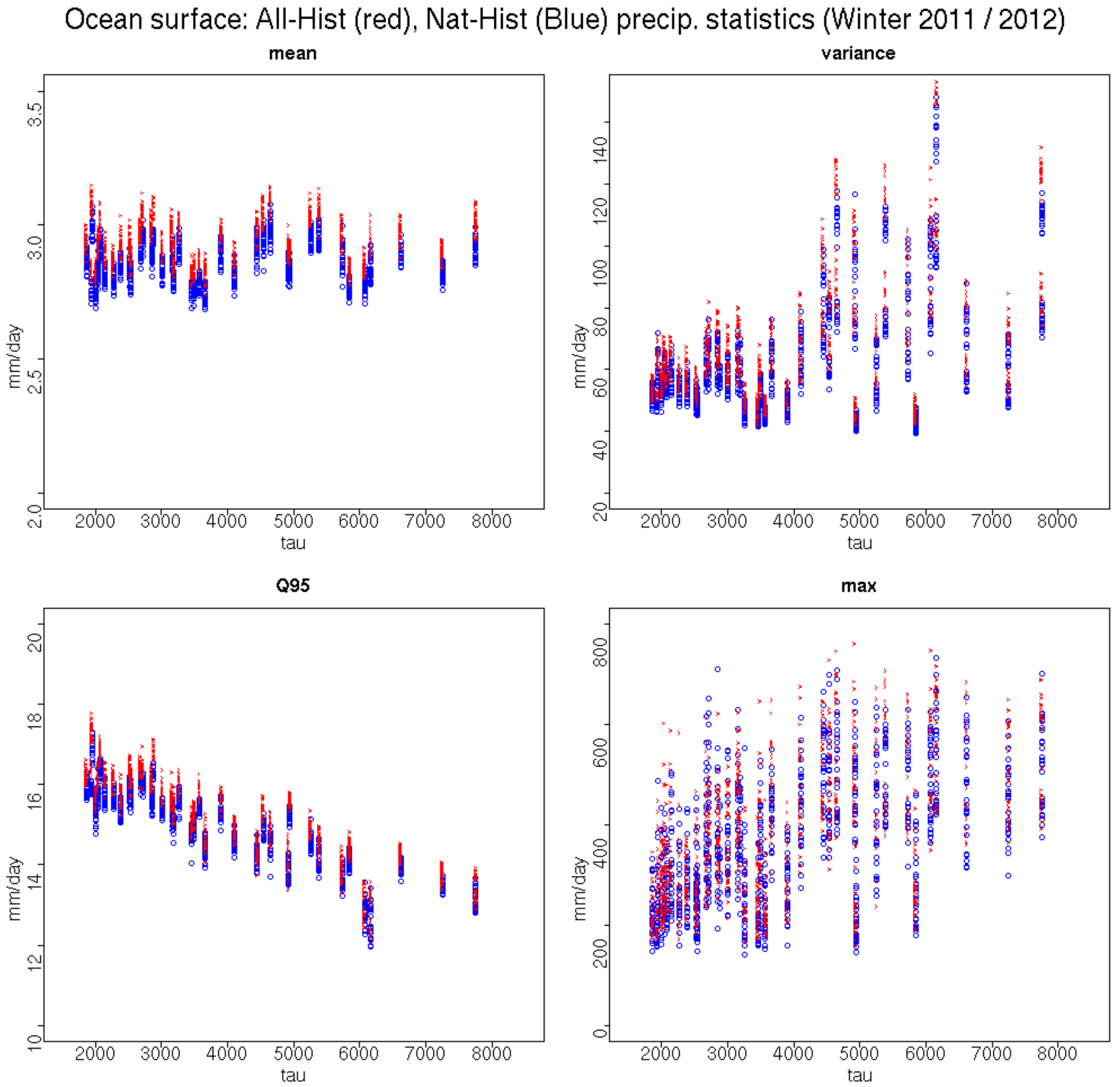
  **4) United States**

- **Blue circles** correspond to Nat-Hist runs, and **red crosses** to All-Hist runs (points may be difficult to distinguish apart).
- Observe the variability in structural dependence across the various summary statistics (mean, variance, Q95, max) by geographic area.

# *1) Ocean surface*

Total precip. (mm / day)

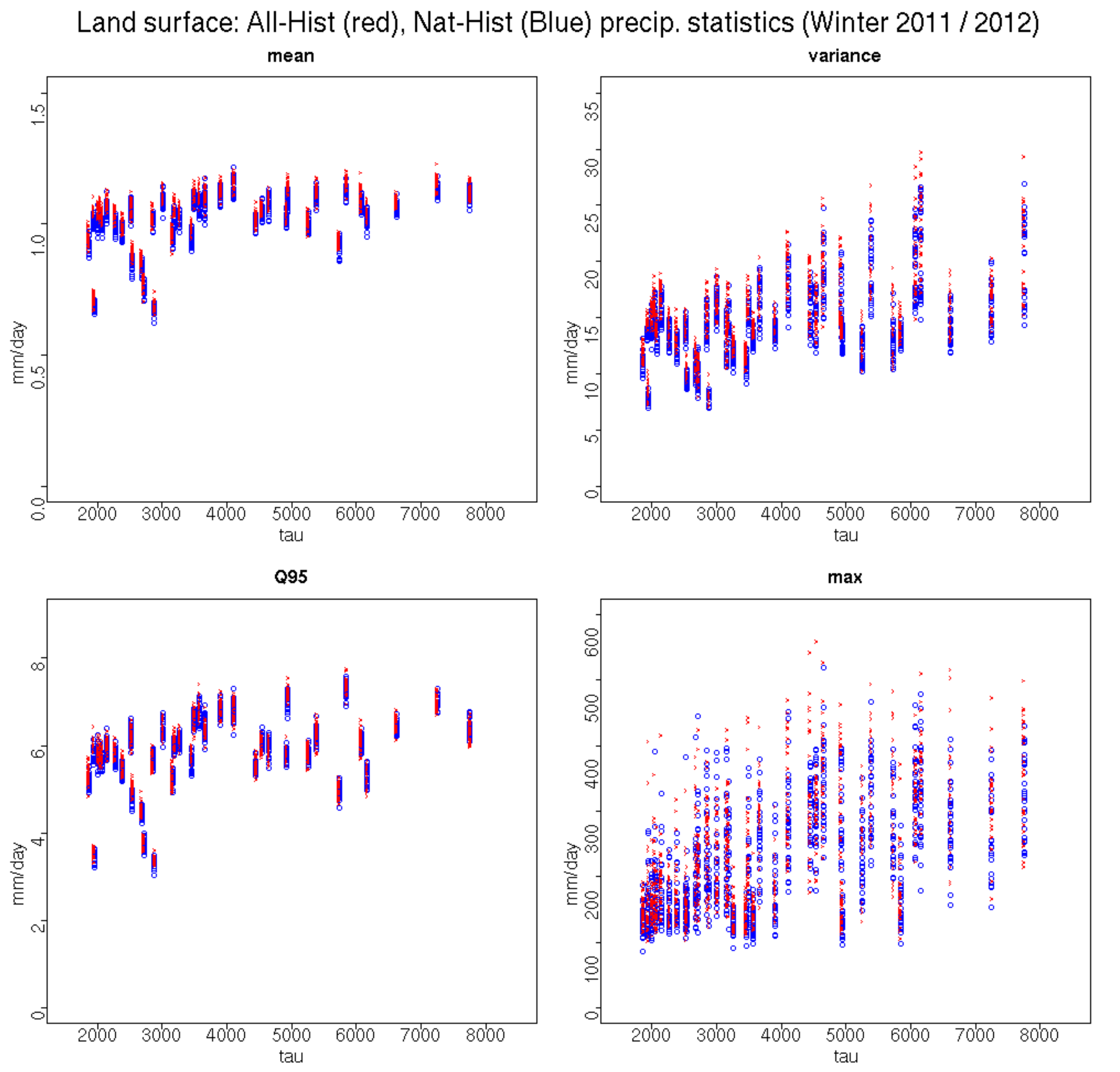
Strong dependence for variance and Q95.



## *2) Land surface*

Total precip. (mm / day)

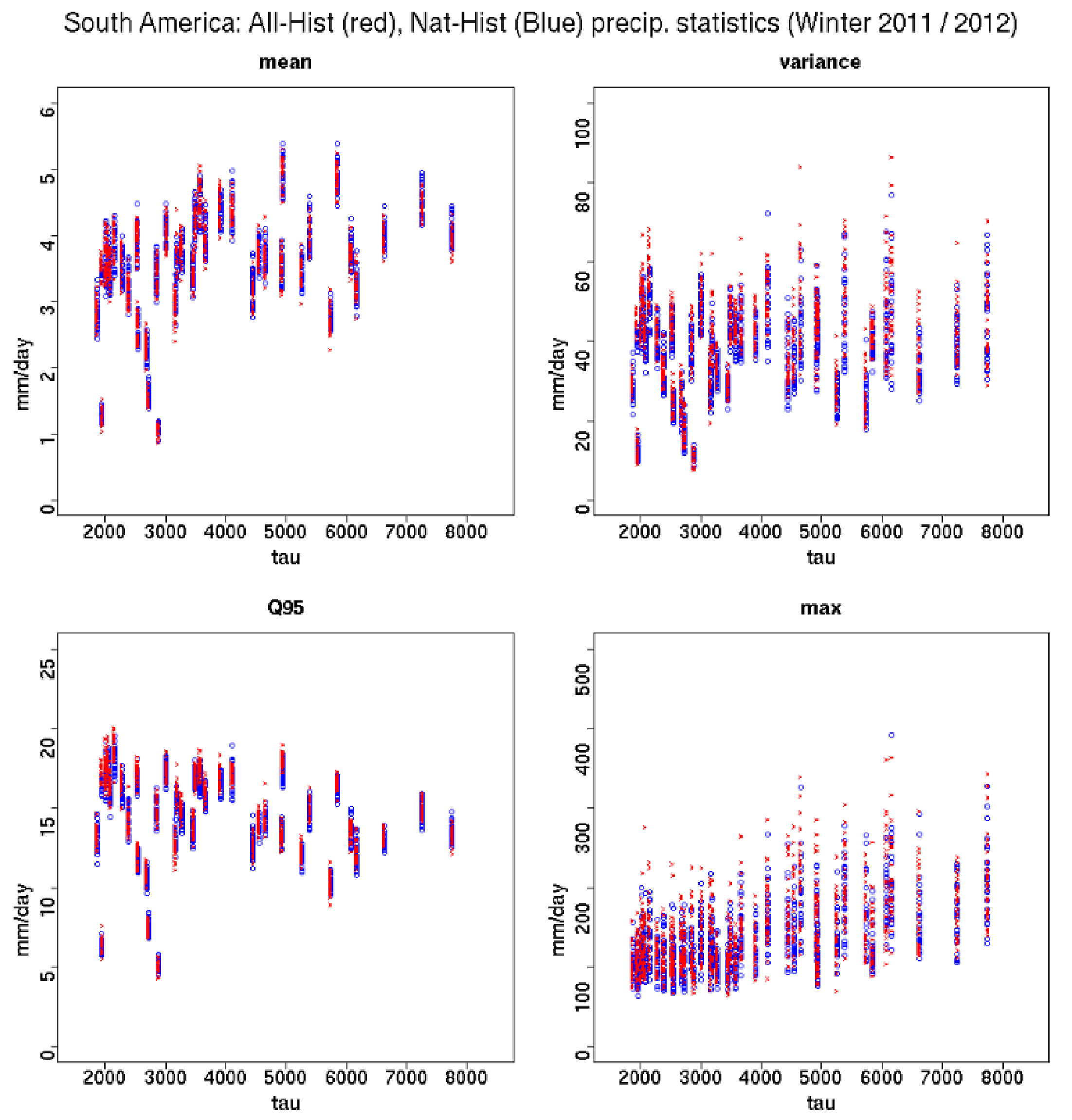
Dependence evident for all summaries.



# *3) South America*

Total precip. (mm / day)

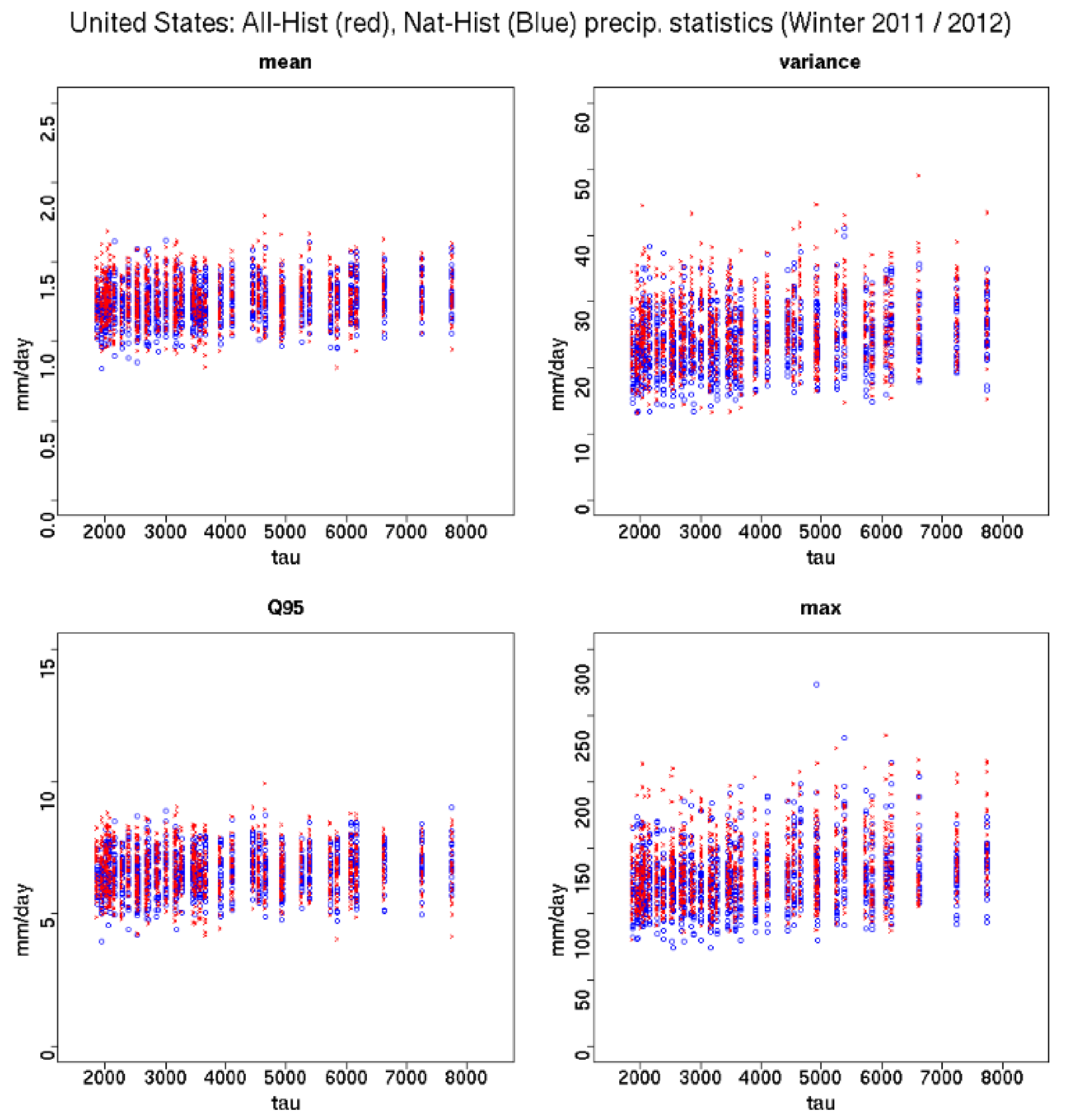
Dependence evident for all summaries.



# *4) United States*

Total precip. (mm / day)

Minimal dependence.



# Emulation and Uncertainty / Sensitivity Analysis

# Initial emulation approach:

- We require sensitivity analysis for “extreme” quantities (e.g. Q95) and Risk Ratios in order to attribute the variance components to different sources:

  Physics parameters

  Internal variability

- Sensitivity analysis requires efficient methods based on surrogate models.
- Gaussian process (GP) based emulators tested in the first instance **(e.g. Oakely & O'Hagan, 2004)**.
- Emulation “marginally” by grid cell (no geospatial dependence) fitted “in bulk” using maximum *a posteriori* estimation for emulator hyperparameters.
- R packages used “*DiceKriging*”.

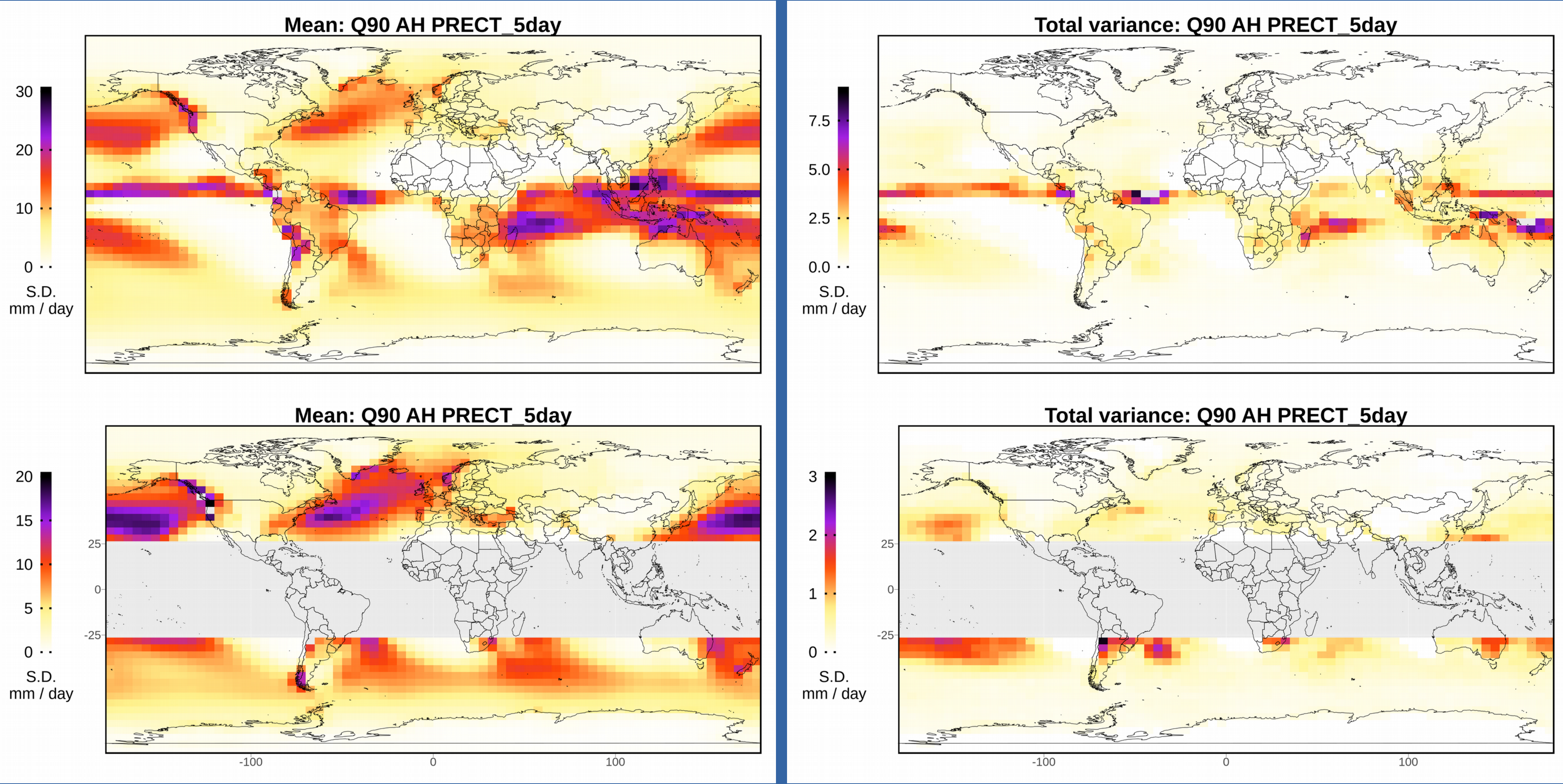
# Results: Global mean and variance Q90 and Q999 PRECT

- With emulators fitted and validated, ensemble mean and total variance is computed at each grid cell. Physics parameter values are sampled from a uniform distribution spanning 90% of design coverage (0.05 to 0.95 parameter range)
- The following slides (1-2) show for DJF (2010 to 2013), All-Hist (AH) only:

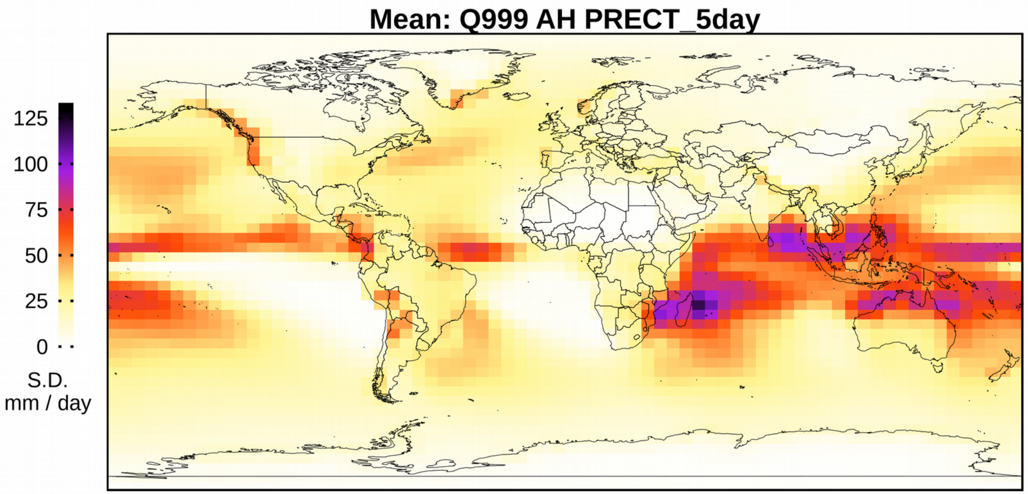
  **1) Q90 PRECT 5-day, mean (left panels) and variance (right panels)**

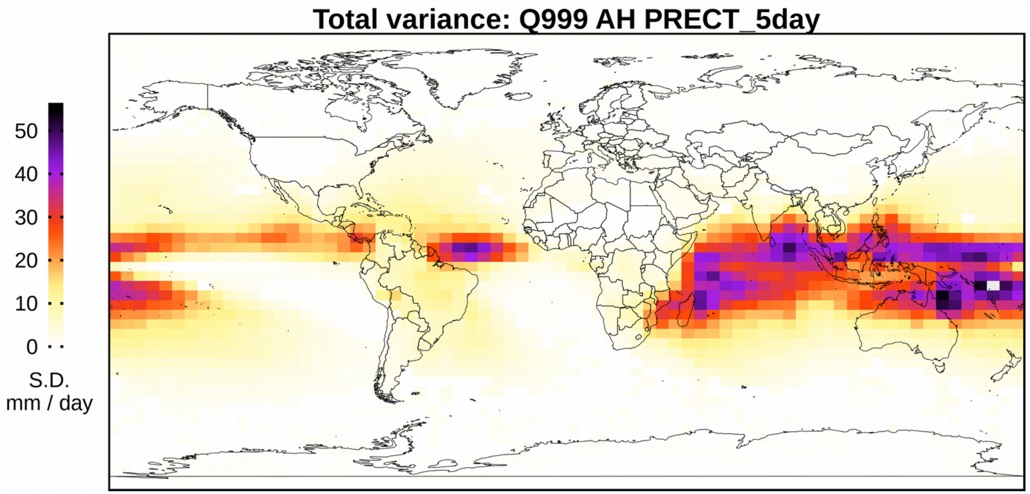
  **2) Q999 PRECT 5-day, mean (left panels) and variance (right panels)**

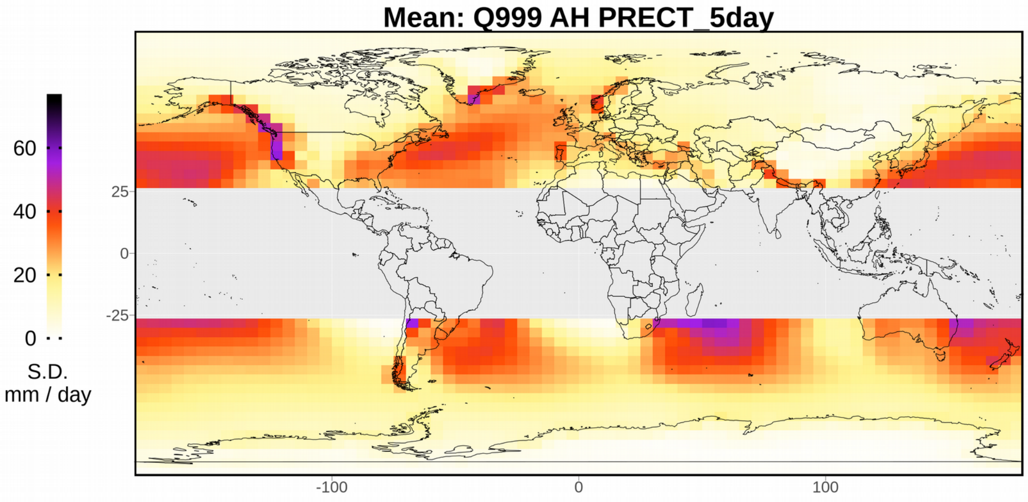
- Lower panels have equatorial region masked, to show structure at higher latitudes.
- Equatorial regions show dramatic variability in precipitation owing to poor representation of atmospheric processes (convection).
- Observe mean / variance ratio is larger for Q90 than Q999.

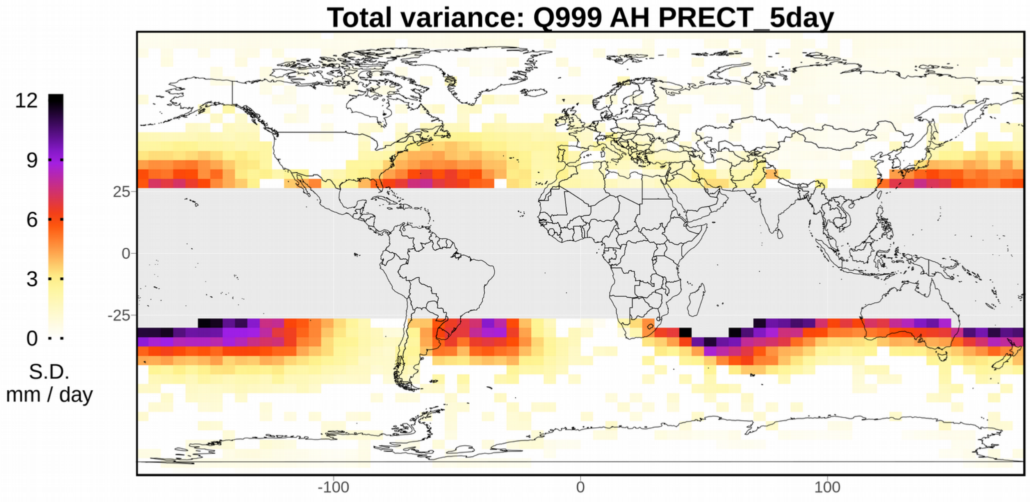


Mean: Q90 AH PRECT_5day
30
20
10
0
S.D.
mm / day
Total variance: Q90 AH PRECT_5day
7.5
5.0
2.5
0.0
S.D.
mm / day
Mean: Q90 AH PRECT_5day
20
15
10
5
0
S.D.
mm / day
25
0
-25
-100
0
100
Total variance: Q90 AH PRECT_5day
3
2
1
0
S.D.
mm / day
25
0
-25
-100
0
100


Mean: Q999 AH PRECT_5day
125
100
75
50
25
0
S.D.
mm / day


Total variance: Q999 AH PRECT_5day
50
40
30
20
10
0
S.D.
mm / day


Mean: Q999 AH PRECT_5day
60
40
20
0
S.D.
mm / day
25
0
-25
-100
0
100


Total variance: Q999 AH PRECT_5day
12
9
6
3
0
S.D.
mm / day
25
0
-25
-100
0
100

# Sensitivity analysis: Global results

# Sensitivity analysis details:

- We require sensitivity analysis for “extreme” quantities (e.g. Q90, Q99, Q99.9) and corresponding Risk Ratios based upon All-Hist and Nat-Hist ensembles, in order to attribute the variance components to different sources:

  Internal variability, physics parameters, climate scenario (for RR), ...

- R packages used “*DiceKriging*” and “*sensitivity*” (sobolGP).

- **Many questions remain:**

  Can we emulate accurately into the extremes (e.g. Q90, Q999)?

  Can the GP model adequately represent internal variability? (Not yet examined in detail)

  Can we identify limits on detection of sensitivity of (e.g. Q90) to physics parameters, given internal variability? (E.g. how many simulated years do we need).

  Can we

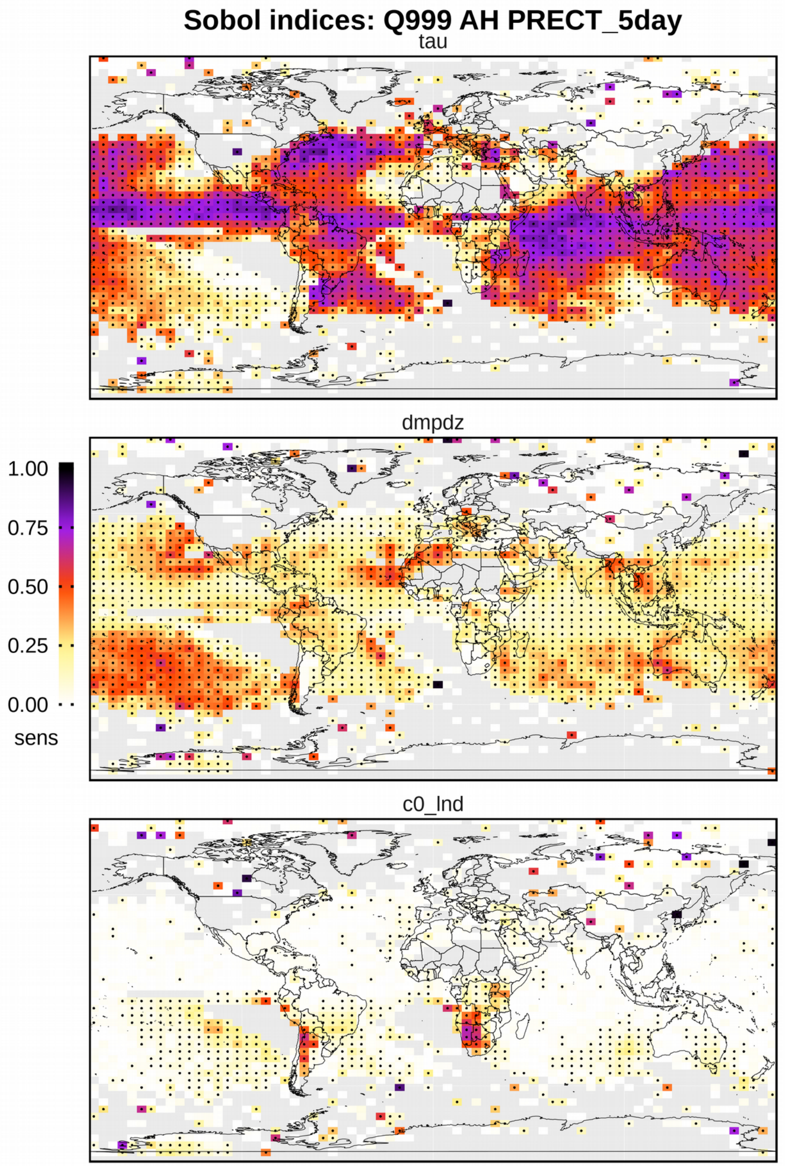
# Some sensitivity analysis (can it work?)

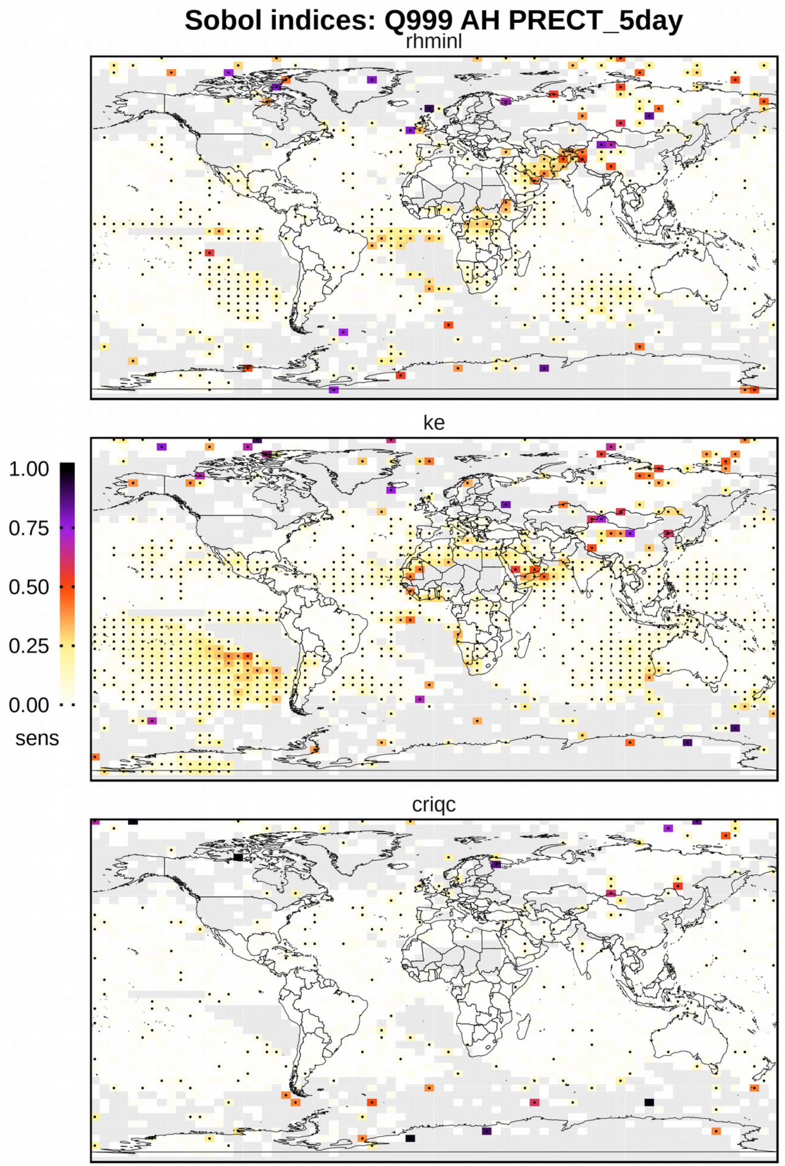
- Following slides (1-2) show Sobol indices: proportion of variance attributable to each parameter (computed for 6 physics parameters) for:

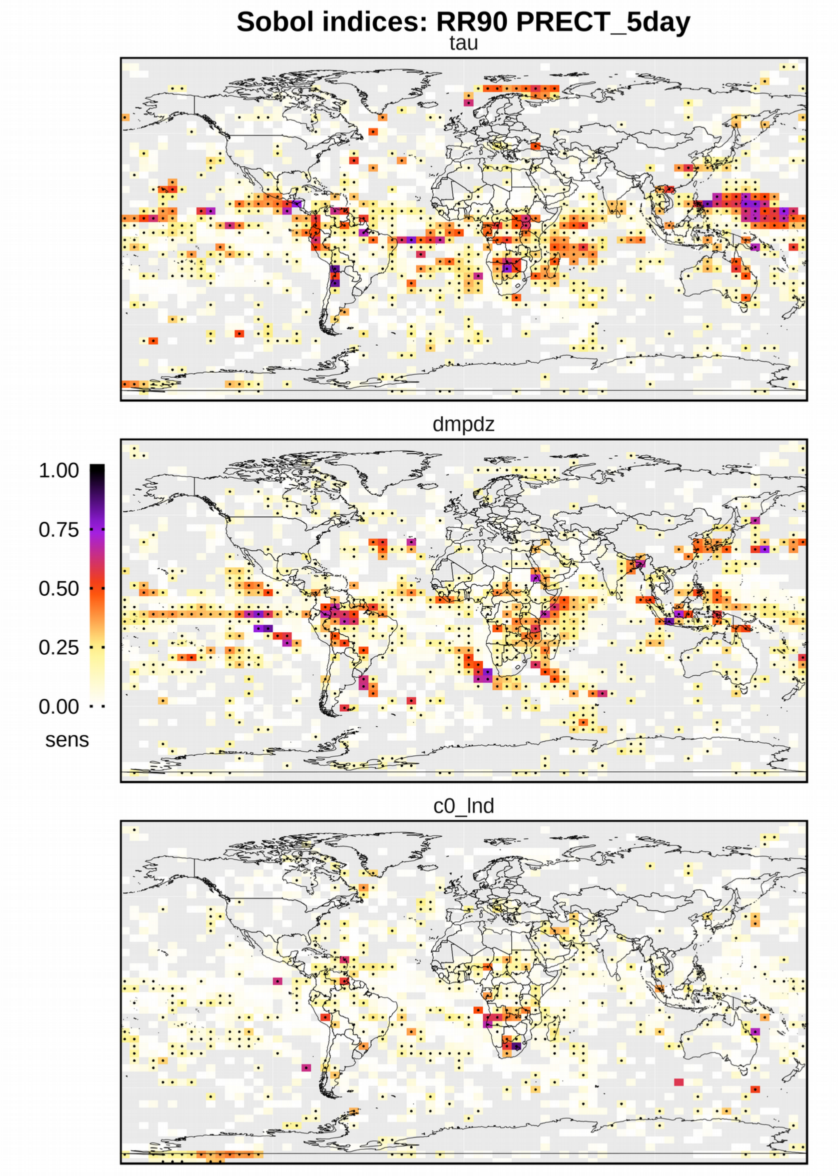
  **1) Q99.9 PRECT 5-day DJF (2010 to 2013), All-Hist (AH) only**

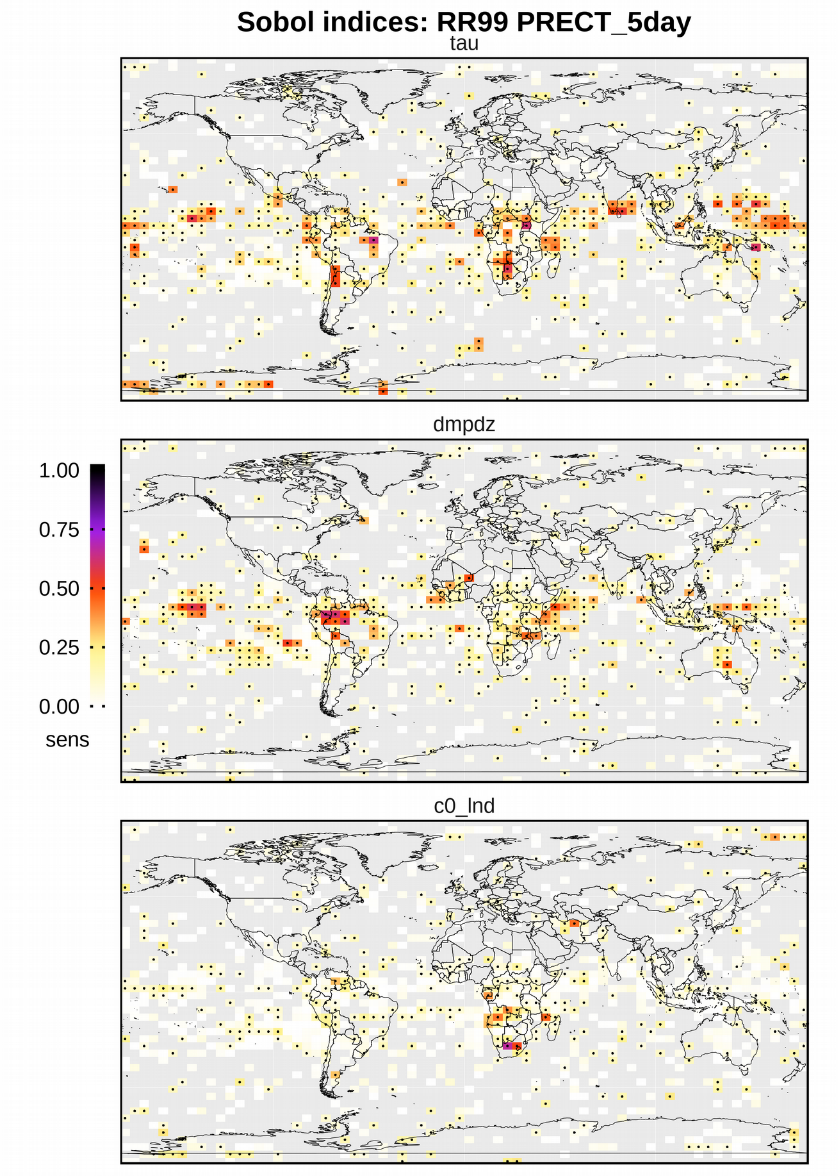
  **2) RR90 (left panel) RR99 (right panel) PRECT 5-day DJF (2010 to 2013)**

- Where RRX is the risk ratio estimated by computing the X %ile from the Nat-Hist simulations, couting daily exceedences in the Nat-Hist and All-Hist, and taking the ratio.
- GP Emulators fitted to QX and RRx, and *sensitivity::SobolGP* R package applied, with uniform prior over 90% (0.05 to 0.95) of parameter value range.
- Grid cells with dots denote statistical significance of Sobol indices.
- Somewhat expensive at scale (1000s processors x few hours).
- Observe that risk ratio (RR90) appears to show sensitivity to perturbed physics in some tropical and equatorial regions, but total variance (not shown) is quite small.


Sobol indices: Q999 AH PRECT_5day
tau
dmpdz
c0_lnd
1.00
0.75
0.50
0.25
0.00
sens


Sobol indices: Q999 AH PRECT_5day
rhminl
ke
criqc
1.00
0.75
0.50
0.25
0.00
sens

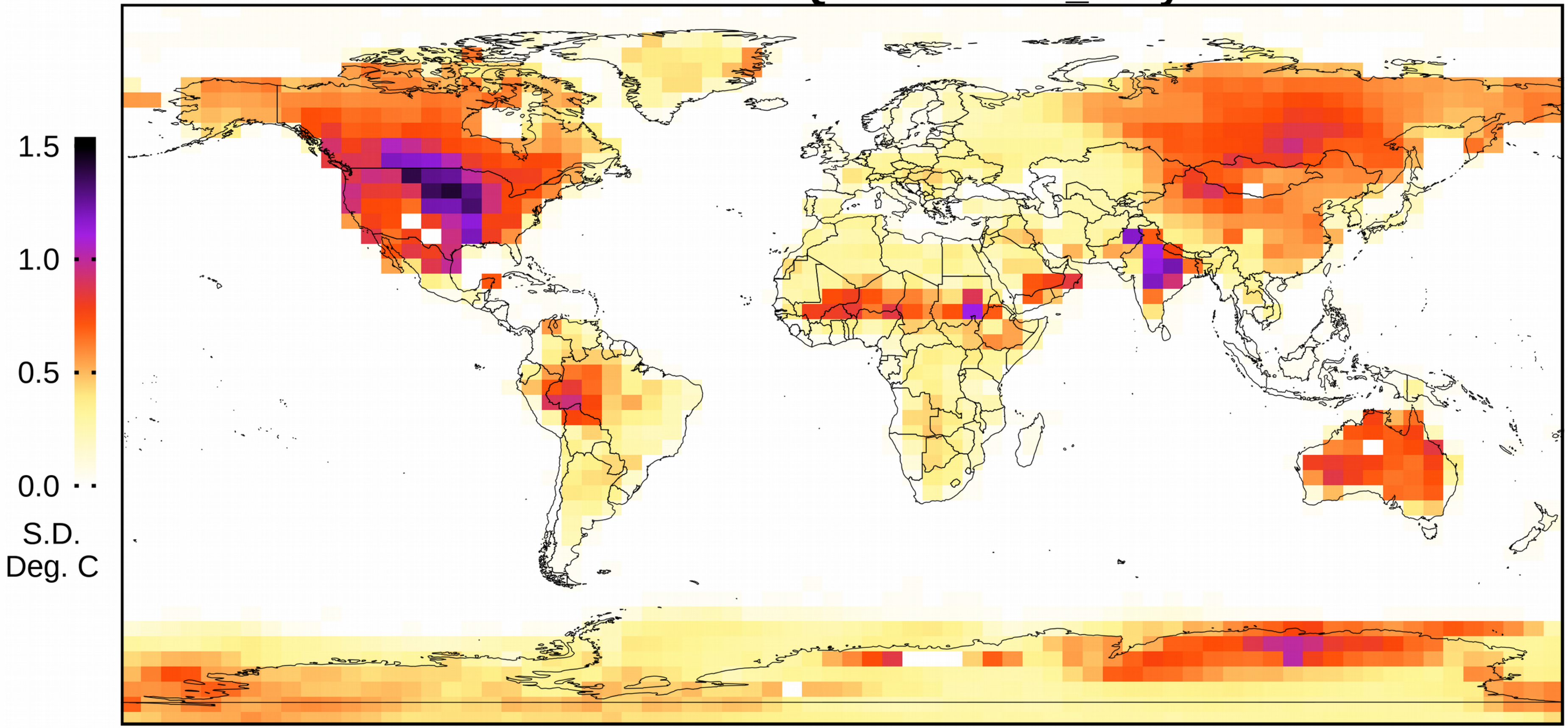

Sobol indices: RR90 PRECT_5day
tau
dmpdz
c0_lnd
1.00
0.75
0.50
0.25
0.00
sens

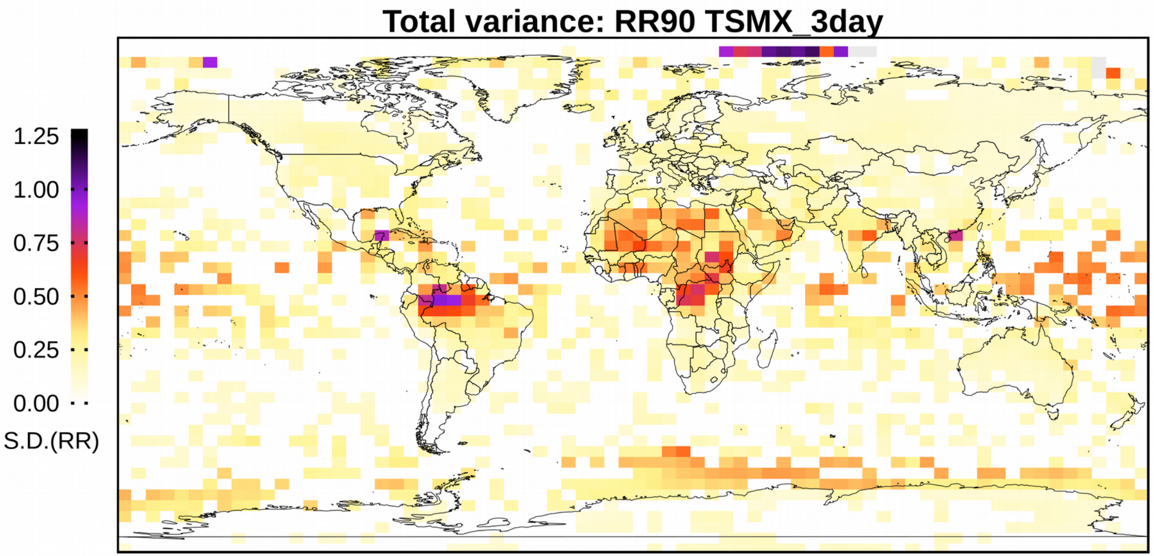

Sobol indices: RR99 PRECT_5day
tau
dmpdz
c0_lnd
1.00
0.75
0.50
0.25
0.00
sens

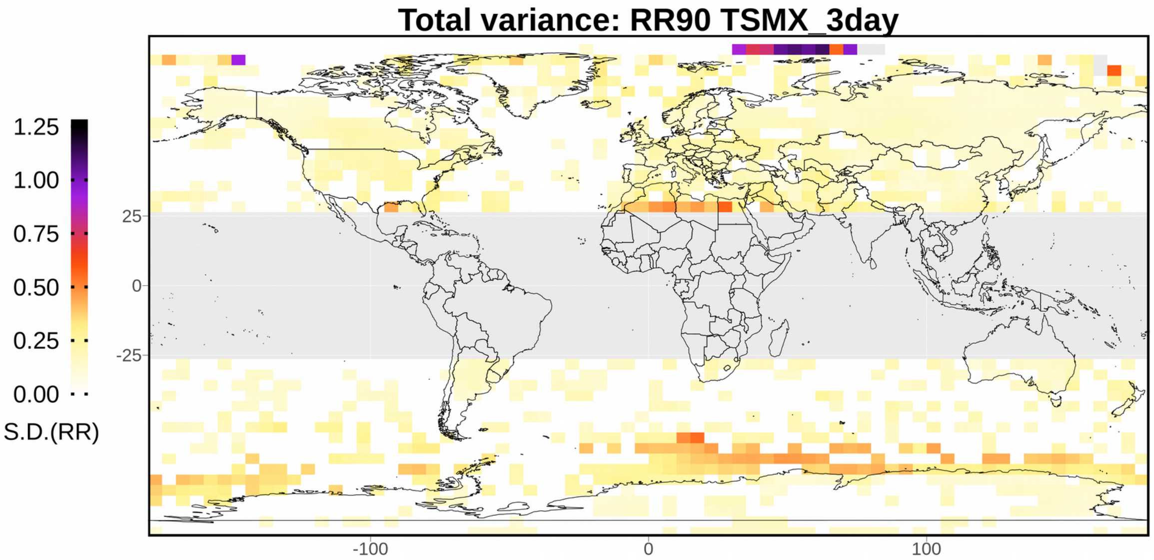
# What about temperature (or other variables)?

- **Not ready yet!**
- Following slides (1-2) show:
    1) **Q90 TSMX 3-day (surface temperature) JJA (2010 to 2013), All-Hist (AH) only**
    2) **Total variance RR90 TSMX 3-day (top panel), equator masked (bottom panel)**
- The risk ratio appears to be sensitive to perturbed physics in equatorial / tropical regions, although total variance is quite small.

# Total variance: Q90 AH TSMX_3day




Total variance: RR90 TSMX_3day
1.25
1.00
0.75
0.50
0.25
0.00
S.D.(RR)


Total variance: RR90 TSMX_3day
1.25
1.00
0.75
0.50
0.25
0.00
S.D.(RR)
25
0
-25
-100
0
100

# Conclusions (so far)

**Ensemble data is available to access! Please contact me with ideas or requests!**

- Initial application of readily deployable tools (GP emulators and sensitivity analysis methods) shows promise.

- Sensitivity of a risk ratio to perturbed physics implies interaction between climate scenario and perturbed physics. Preliminary results suggests that, for CESM, this may be the case in convectively active regions (tropics / ocean). Therefore, attribution statements based on derived risk ratios could be affected in such regions. (Note however that representation of sub-grid scale convective processes is poor in then tropics.)

- **Many open questions remain:**

  Detailed examination of how emulators capture internal variability is required.

  Emulator refinement  “borrowing strength” through spatial modeling?

  **Can other machine learning / AI methods be applied effectively?**

  Focus on specific attribution events / studies to establish limits of detectability of effects of perturbed physics ( -> augmentation of design or additional replicates).

# References


- Qian *et al.* (2015) “Parametric sensitivity analysis of precipitation at global and local scales in the Community Atmosphere Model CAM5” *JAMES*
- Williamson, D. (2015) “Exploratory ensemble designs for environmental models using k-extended Latin Hypercubes” *Environmetrics*
- Angelil *et al.* (2017) “An Independent Assessment of Anthropogenic Attribution Statements for Recent Extreme Temperature and Rainfall Events” *J. Clim.*
- Kooperman *et al.* (2018) “Rainfall from resolved rather than parameterized processes better represents the present-day and climate change response of moderate rates in the community atmosphere model.” *JAMES* https://doi.org/10.1002/2017MS001188
- Stone *et al.* (2019), “Experiment design of the International CLIVAR C20C+ Detection and Attribution project” *WACE* https://doi.org/10.1016/j.wace.2019.100206
- Oakley & O'Hagan (2004) “Probabilistic sensitivity analysis of complex models: a Bayesian approach” *J. R. Statist. Soc. B*

# Basic emulator formulation

With some *Simulator* producing output $y$ for given input $\mathbf{x}$,

$$y = \eta(\mathbf{x})$$

But we do not know $y$ for all $\mathbf{x}$. Adopting a Gaussian process model for $\eta(\cdot)$ we have,

$$\eta(\mathbf{x})|\beta, \sigma^2 \sim \mathcal{GP}(\mu(\mathbf{x}), \sigma^2 c(\mathbf{x}, \mathbf{x}'))$$

with mean and correlation functions,

$$\begin{aligned} \mu(\mathbf{x}) &= h(\mathbf{x})^{\mathrm{T}}\beta \\ c(\mathbf{x}, \mathbf{x}') &= exp\left(-(\mathbf{x} - \mathbf{x}')^{\mathrm{T}} B(\mathbf{x} - \mathbf{x}')\right) \end{aligned}$$

where $h(\cdot)$ are basis functions, $\beta$ is a vector coefficients and $B$ is a diagonal matrix of length scale parameters.

Using a Bayesian approach, prior distributions are specified for $\beta, \sigma^2$ which are typically weak,

$$p(\beta, \sigma^2) \propto \sigma^{-2}$$

In order to update our prior model with known Simulator output, we can make "observations" $y$ of the Simulator, at $n$ specified input points $\mathrm{x}_1 \ldots \mathrm{x}_n$ according to a design such that,

$$\mathbf{y}^{\mathrm{T}} = [\eta(\mathrm{x}_1), \eta(\mathrm{x}_2), \ldots, \eta(\mathrm{x}_n)]$$

According to our model then, the data conditional on $\sigma^2$ and $\beta$ is,

$$\mathbf{y}|\beta, \sigma^2 \sim \mathcal{GP}(H\beta, \sigma^2\mathbf{A})$$

where,

$$H^{\mathrm{T}} = [h(\mathrm{x}_1), h(\mathrm{x}_2), \ldots, h(\mathrm{x}_n)]$$

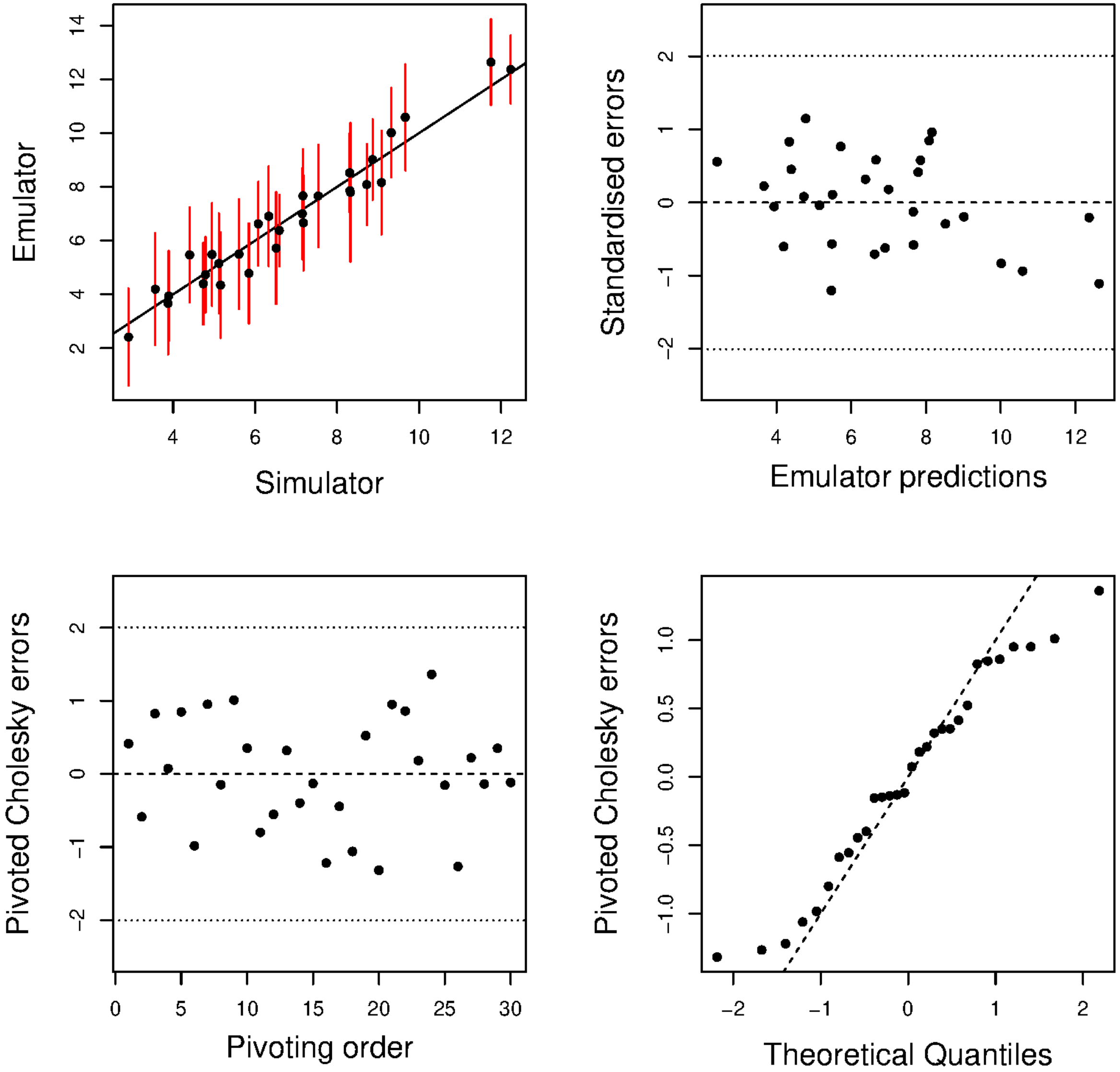
$$\mathbf{A} = \begin{pmatrix} 1 & c(\mathrm{x}_1, \mathrm{x}_2) & \ldots & c(\mathrm{x}_1, \mathrm{x}_n) \\ c(\mathrm{x}_2, \mathrm{x}_1) & 1 & & \vdots \\ \vdots & & \ddots & \\ c(\mathrm{x}_n, \mathrm{x}_1) & \ldots & & 1 \end{pmatrix}$$

The model is updated by conditioning our prior model on the simulator output giving the following,

$$\eta(\mathbf{x})|\beta, \sigma^2, \mathbf{y} \sim \mathcal{GP}(m^*(\mathbf{x}), \sigma^2 c^*(\mathbf{x}, \mathbf{x}'))$$

where,

$$
\begin{aligned}
m^*(\mathbf{x}) &= \mathbf{h}(\mathbf{x})^{\mathrm{T}}\beta + \mathbf{t}(\mathbf{x})^{\mathrm{T}}\mathbf{A}^{-1}(\mathbf{y} - H\beta) \\
c^*(\mathbf{x}, \mathbf{x}') &= c(\mathbf{x}, \mathbf{x}') - \mathbf{t}(\mathbf{x})^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{t}(\mathbf{x}') \\
\mathbf{t}(\mathbf{x})^{\mathrm{T}} &= [c(\mathbf{x}, \mathbf{x}_1), \ldots, c(\mathbf{x}, \mathbf{x}_n)]
\end{aligned}
$$


Emulator
Simulator
Standardised errors
Emulator predictions
Pivoted Cholesky errors
Pivoting order
Pivoted Cholesky errors
Theoretical Quantiles