



# Identifying and emulating hazard metrics


November 5, 2019

**Claudia Tebaldi**  
JGCRI



PNNL is operated by Battelle for the U.S. Department of Energy





**We want to incorporate metrics of hazards, (exposure, impacts) in integrated modeling**

- 1) Identify **impact relevant metrics of hazards** for different sectors (energy, infrastructure, agriculture, ecosystem,...).
- 2) Analyze ESM output deriving **statistics of current and future behavior** for these metrics.
- 3) Emulate this behavior in a **computationally efficient** manner so that it can be represented in integrated modeling, with the ultimate goal of **representing feedbacks from Earth to human systems and vice-versa, i.e., exploring scenarios within which impacts shape future development trajectories.**
- 4) Conduct **sensitivity** studies and **uncertainty** quantification, enabled by computationally efficient methods.

## Some initial exploration and future plans

- Choose ten indices of extreme temperature and precipitation;
- Exploit several initial condition ensembles run by CESM under a range of future emission trajectories;
- Test two alternative emulation methods and determine their performance.

- 
- Implement emulation within the GCAM ecosystem using Hector;
  - Extend emulation to more metrics of hazards;
  - Extend emulation to impact metrics, starting from exposure (population, cultivated areas) and ideally extending all the way to actual impacts (by including vulnerability measures).

## **Five indices of temperature extremes:**

- Frost days
- Growing season length
- Hottest day of the year
- Coldest day of the year
- Warm spells

## **Five indices of precipitation extremes:**

- Max dry spell duration
- Precipitation intensity
- Days with >10mm
- Total precipitation falling during wettest days
- Wettest 5-day period



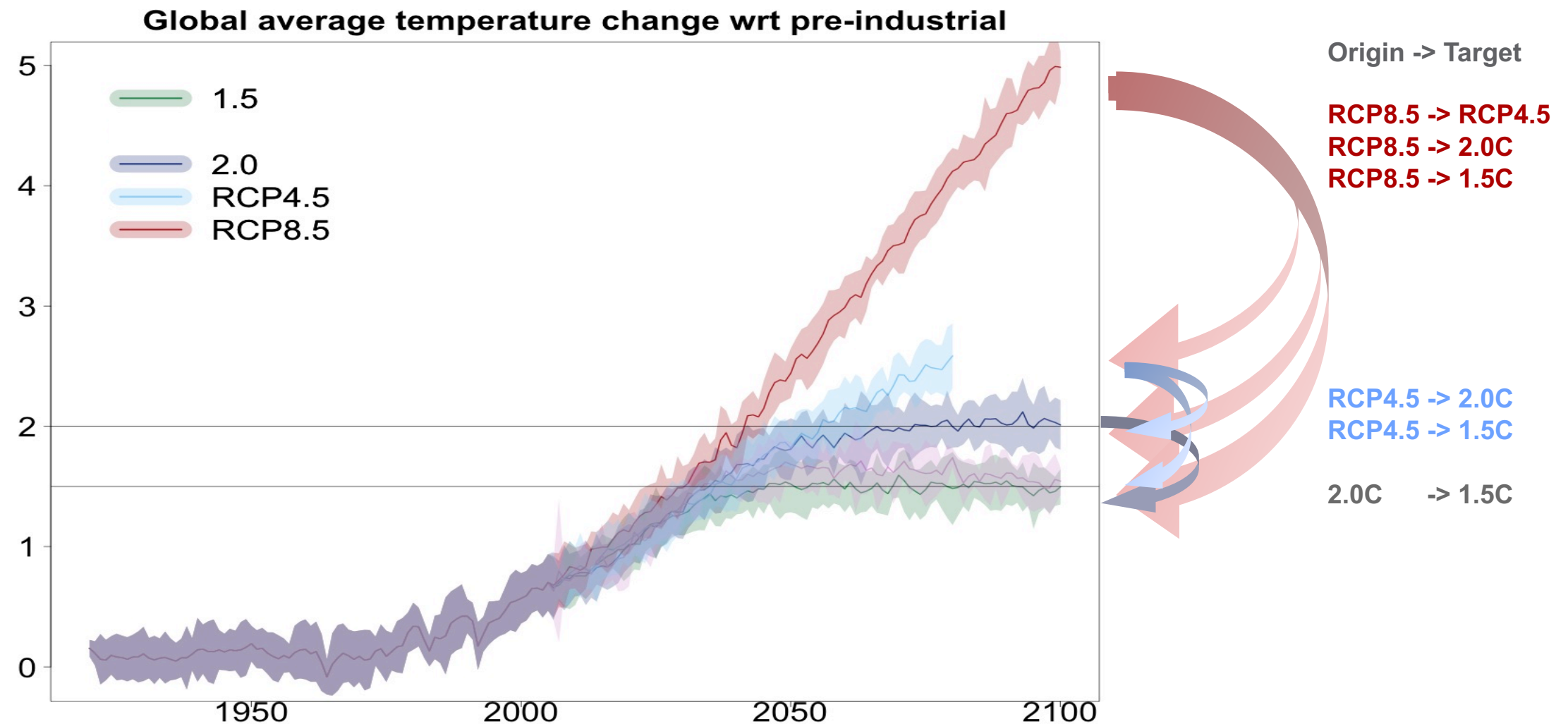
## Pattern Scaling

- Compute **time- and scenario-independent normalized patterns of change** (per degree of global warming) from available ESM simulations.
- Use **global average temperature from simple model to rescale** and thus simulate patterns for new scenarios and time periods.

## Time Shift

- Use **output of ESM simulations at time when global average temperature is the same** as target scenario/time period.

# Origin (training data) and Target (truth) Scenarios



$(Y_1, \dots, Y_{10})$  replicates of the target quantity from  $n$  ensemble members;  
 $(\hat{Y}_1, \dots, \hat{Y}_{10})$  emulated quantities derived from  $n$  ensemble members of the origin scenario;

Two measures of performance:

- How good is the emulation “on average”, i.e., what is the size of the true emulation error  $\sqrt{\langle \bar{Y} - \bar{\hat{Y}} \rangle^2}$
- How does the internal variability of the emulation replicate the internal variability of the target

$$\sqrt{\langle (\hat{Y} - \bar{\hat{Y}})^2 \rangle} \text{ vs } \sqrt{\langle (Y - \bar{Y})^2 \rangle}$$

So we propose the two-dimensional error metric

$$\left( \sqrt{\langle \bar{Y} - \bar{\hat{Y}} \rangle^2} / \sqrt{\langle (Y - \bar{Y})^2 \rangle}; \sqrt{\langle (\hat{Y} - \bar{\hat{Y}})^2 \rangle} / \sqrt{\langle (Y - \bar{Y})^2 \rangle} \right)$$

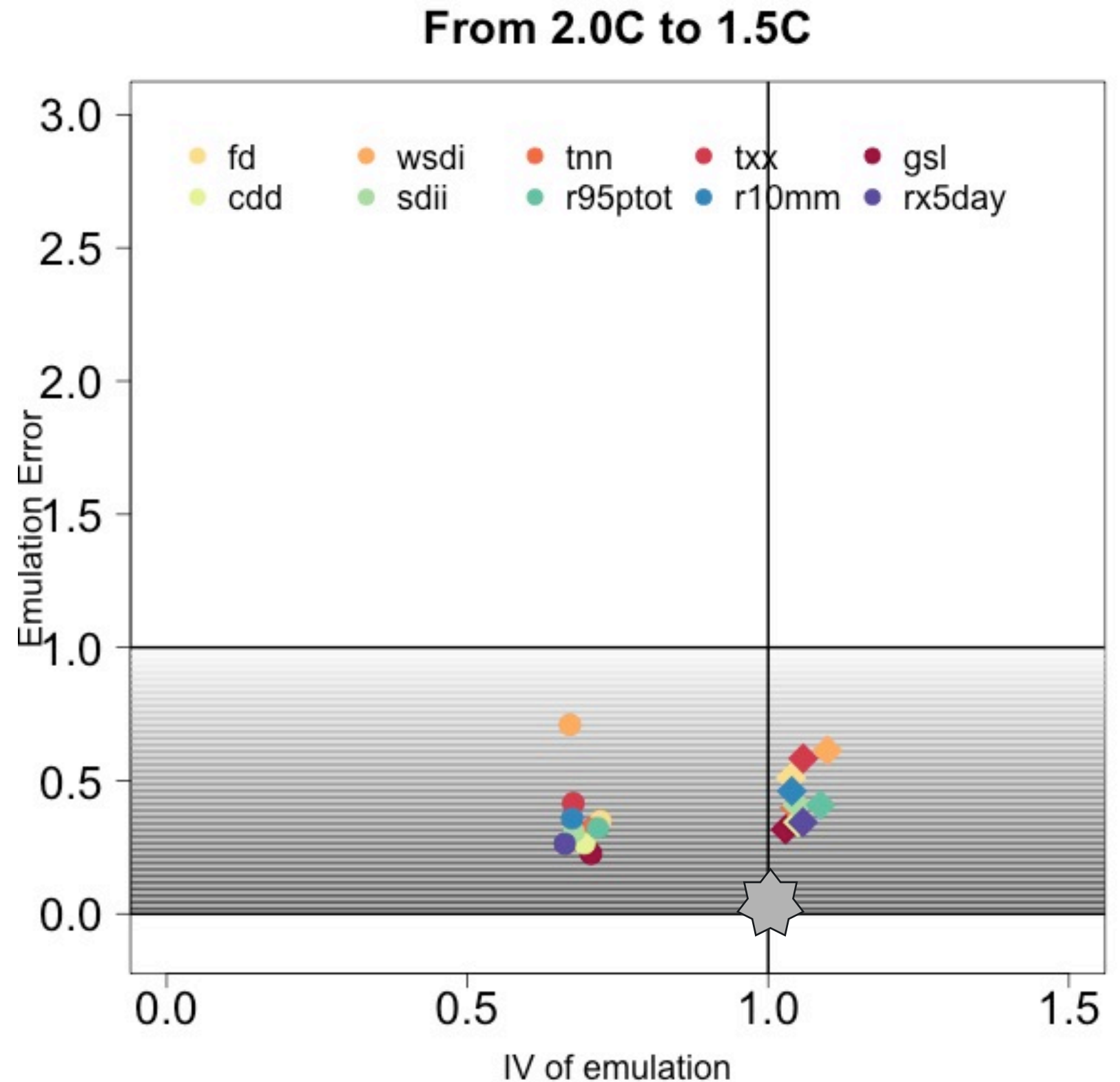
## Components of the Metric of performance:

Emulation Error:

$$\sqrt{\langle \bar{Y} - \hat{\bar{Y}} \rangle^2} / \sqrt{\langle (Y - \bar{Y})^2 \rangle}$$

IV of Emulation:

$$\sqrt{\langle (\hat{Y} - \bar{\hat{Y}})^2 \rangle} / \sqrt{\langle (Y - \bar{Y})^2 \rangle}$$





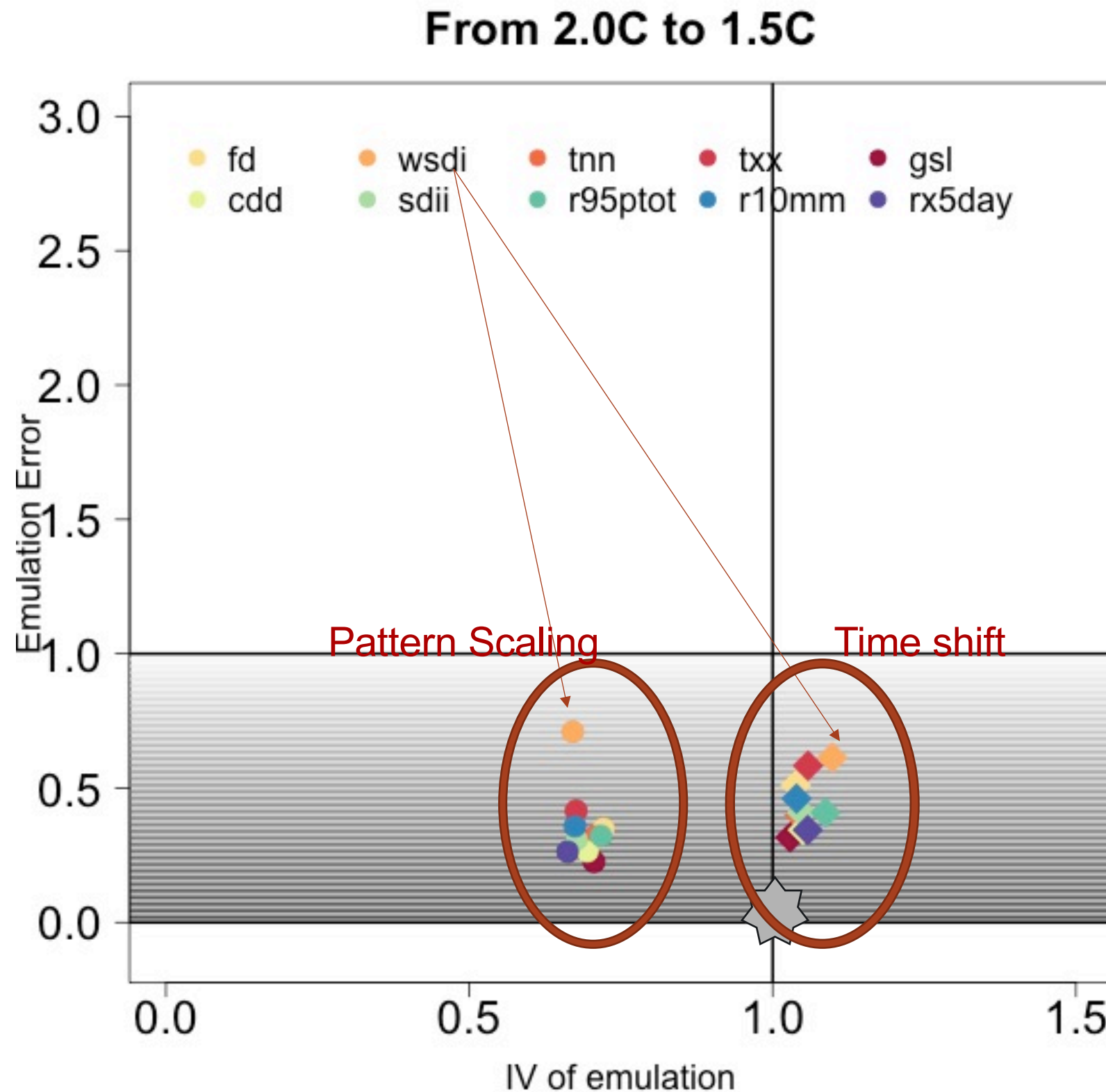
## Components of the Metric of performance:

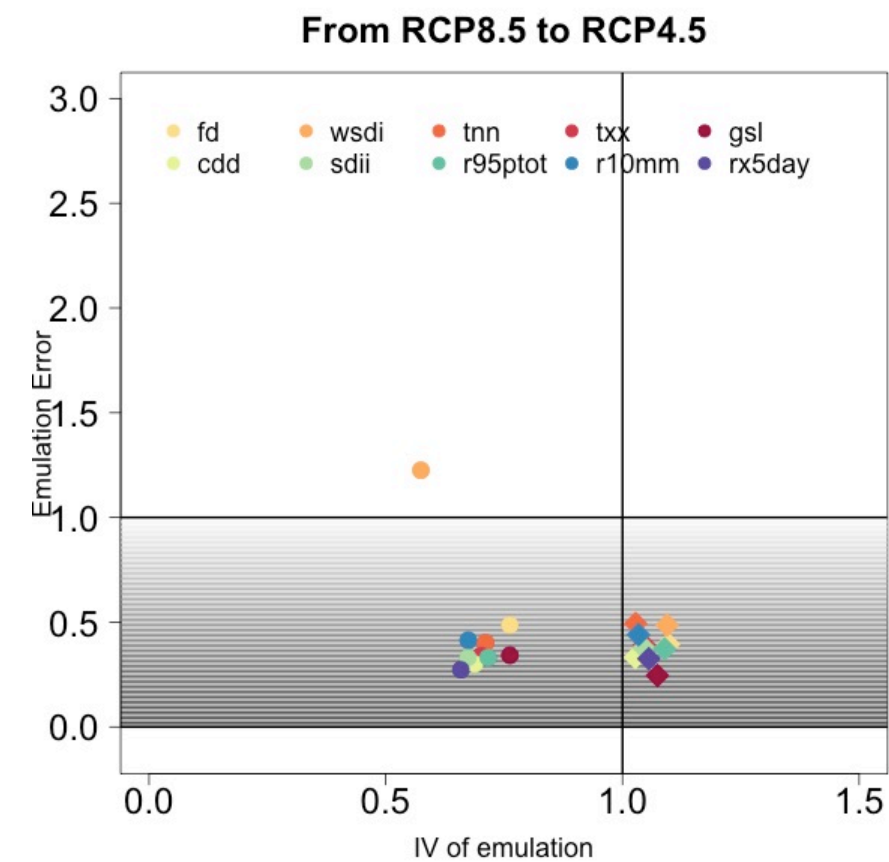
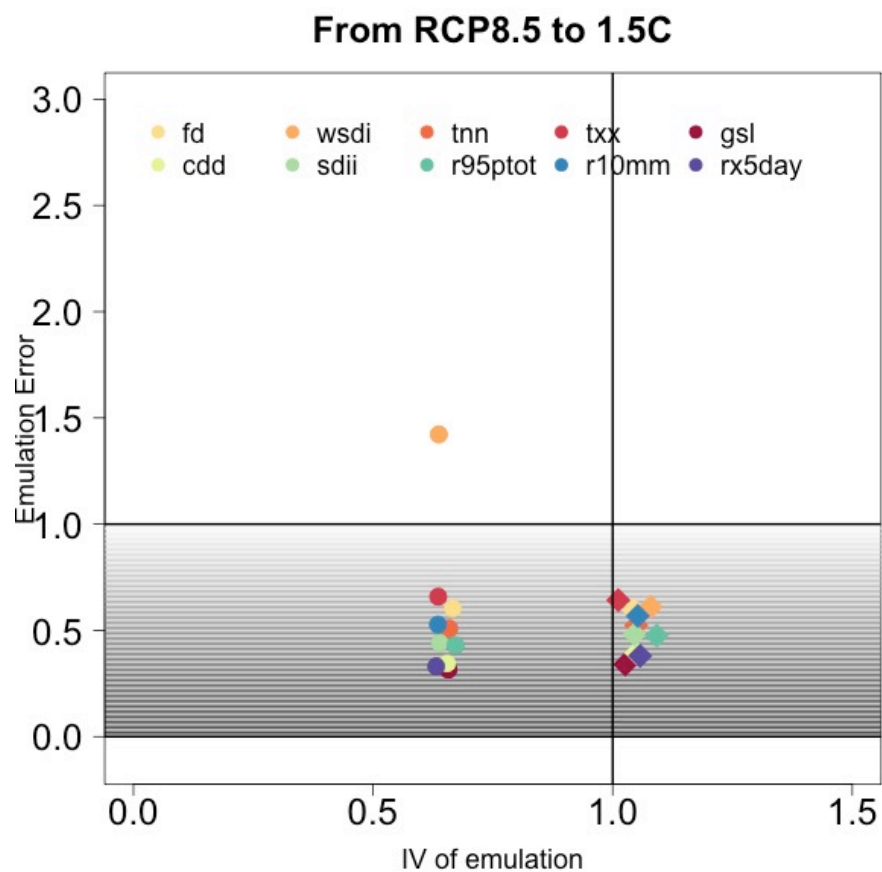
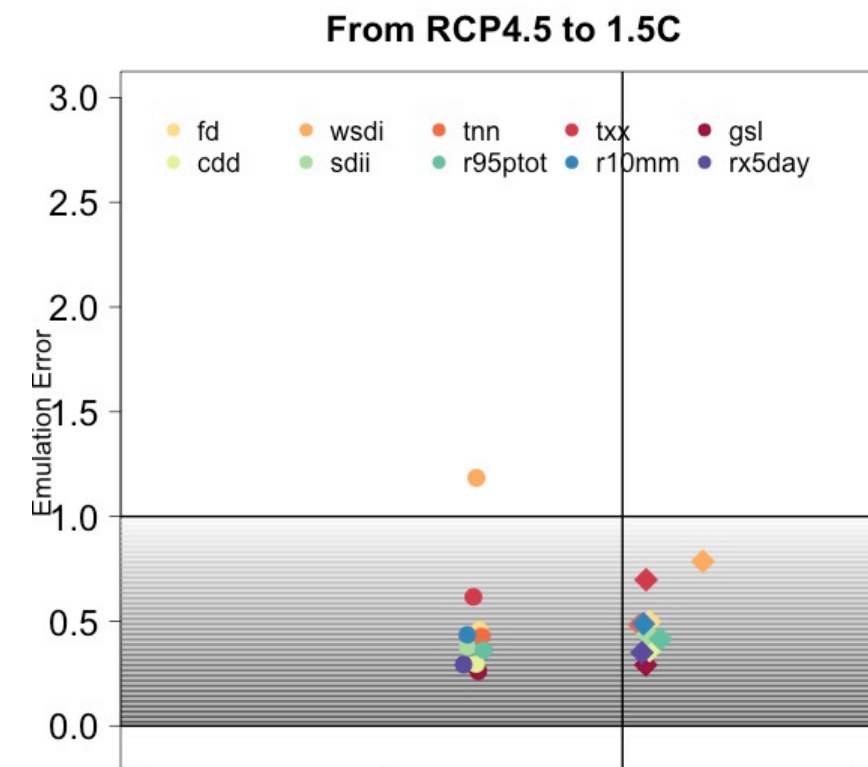
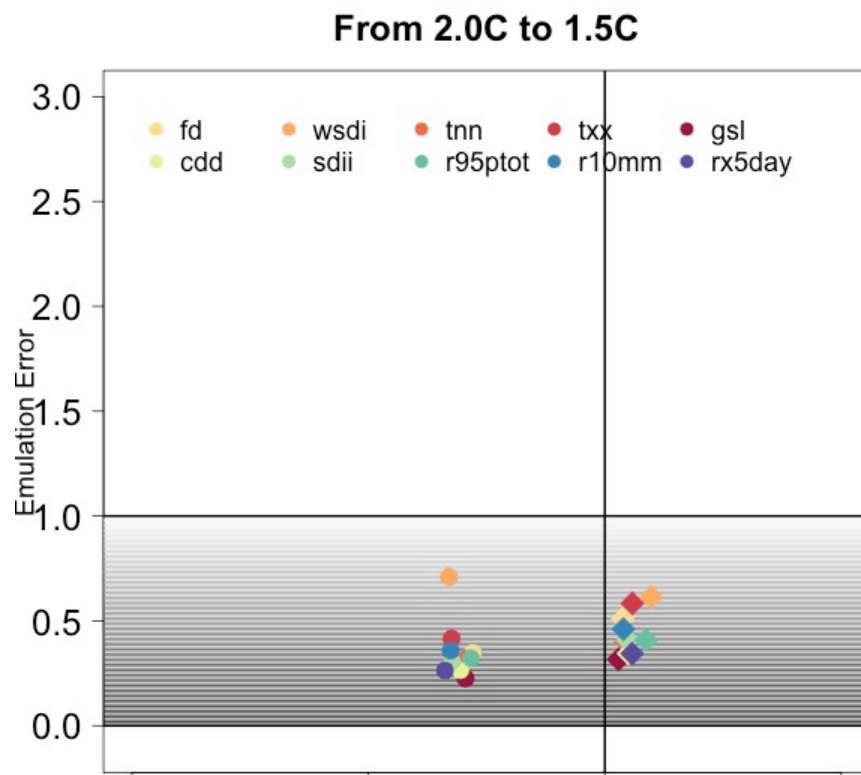
Emulation Error:

$$\sqrt{\langle \bar{Y} - \hat{\bar{Y}} \rangle^2} / \sqrt{\langle (Y - \bar{Y})^2 \rangle}$$

IV of Emulation:

$$\sqrt{\langle (\hat{Y} - \bar{\hat{Y}})^2 \rangle} / \sqrt{\langle (Y - \bar{Y})^2 \rangle}$$





## So far

Nine out of ten indices are well emulated by both methods, with errors that are smaller than the variation expected because of internal variability. This is true for both temperature and precipitation indices.

Pattern Scaling slightly outperforms Time Shift in terms of emulation error, but Time Shift comes closer to replicate the internal variability of the target, which is a desirable feature for uncertainty representation/quantification.

The gap between scenario does not appear to be a large factor in the overall performance of the emulation.

Only one index appears problematic for pattern scaling, hinting at non linearities.

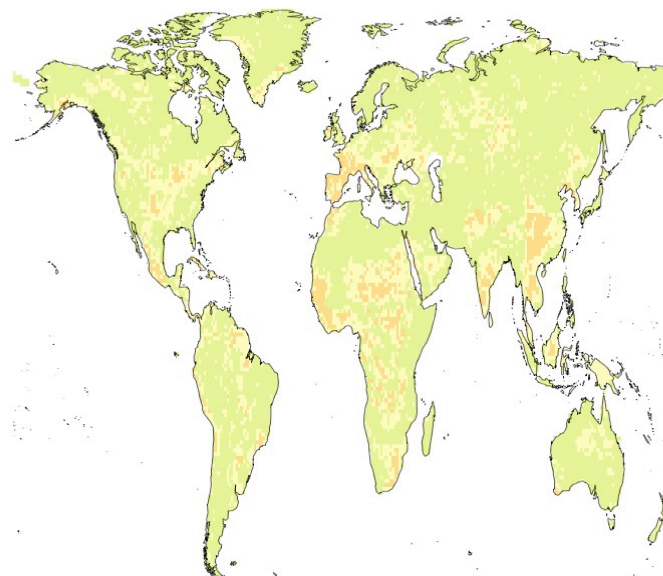


# What happens geographically

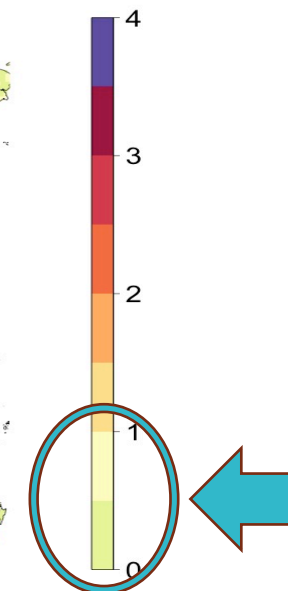
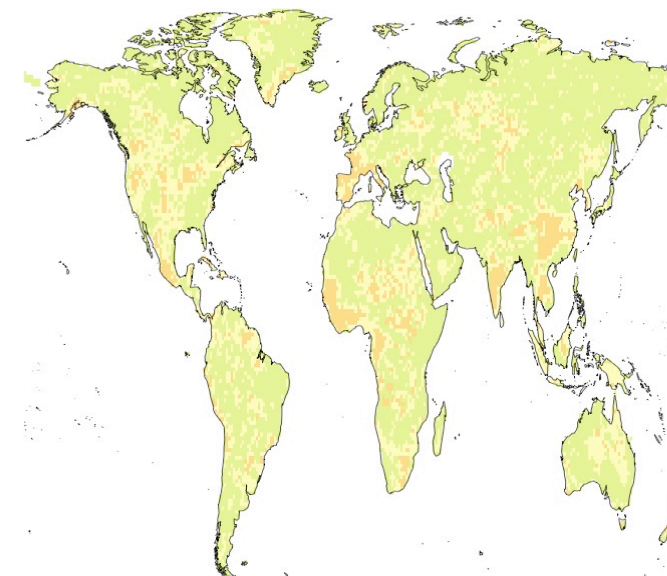
## A good emulation:

- SDII:  
Precipitation Intensity

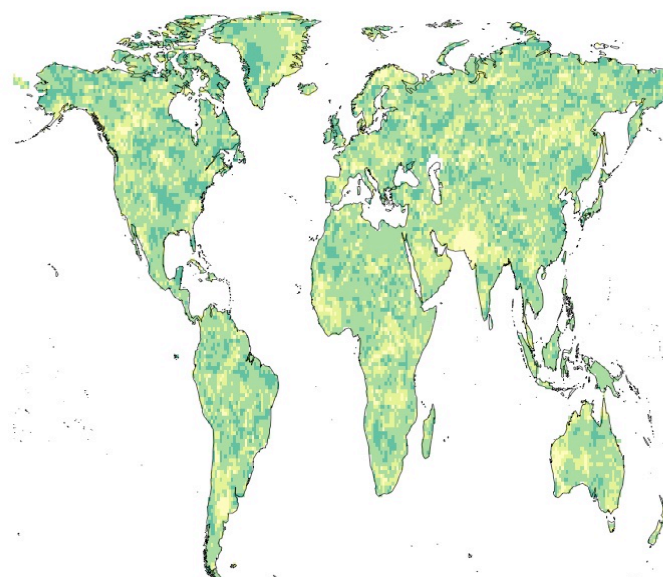
Emulation error



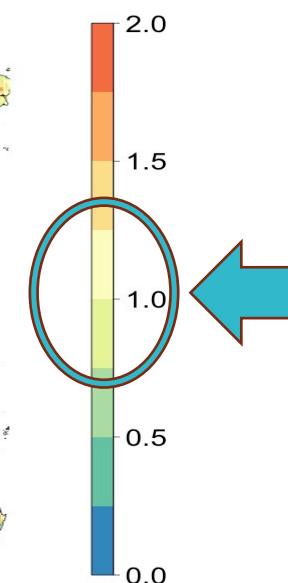
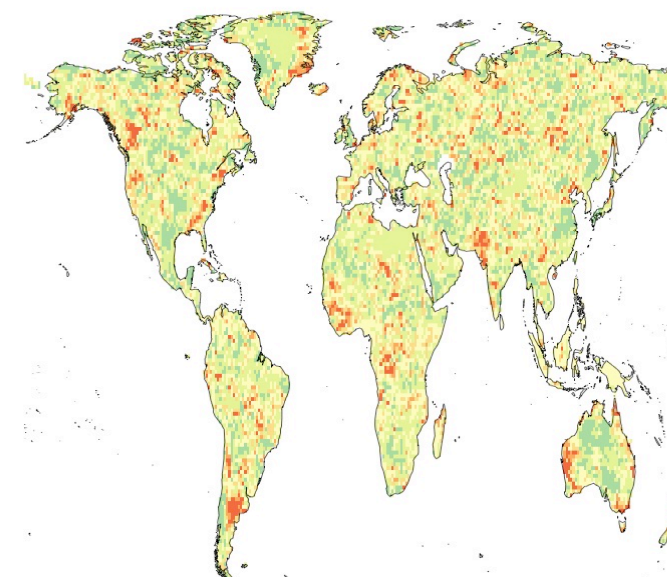
Emulation error



IV of Emulation



IV of Emulation

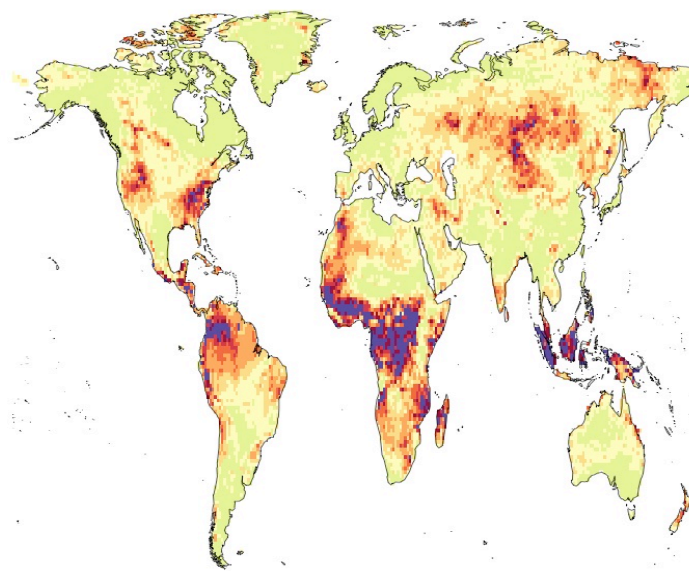


# What happens geographically

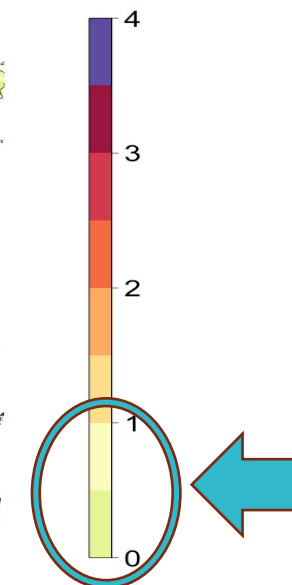
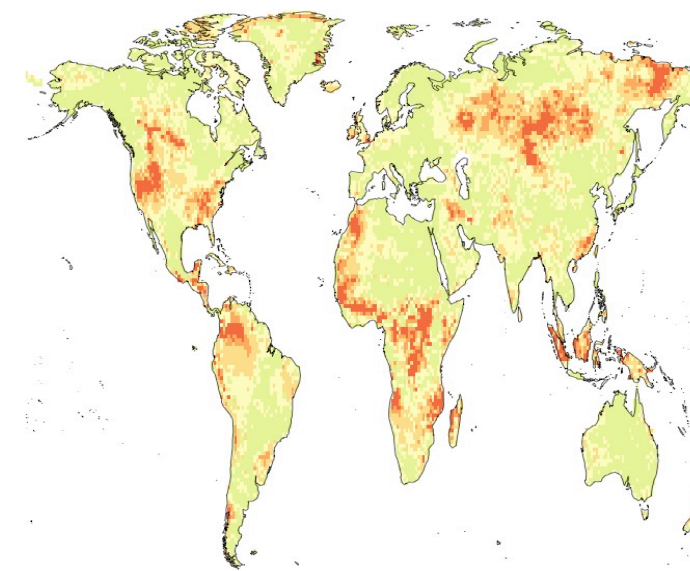
## The bad emulation:

- WSDI:  
Warm spells

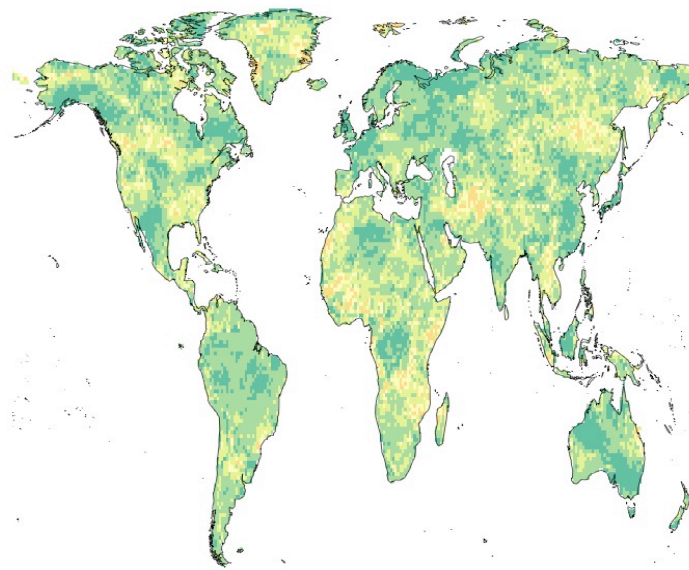
Emulation error



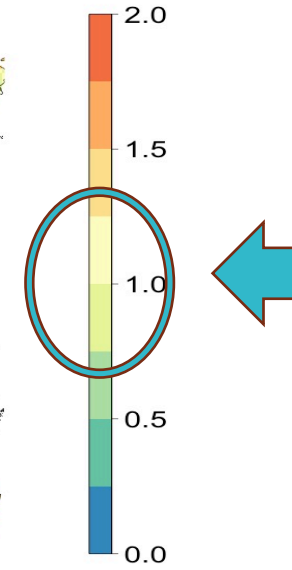
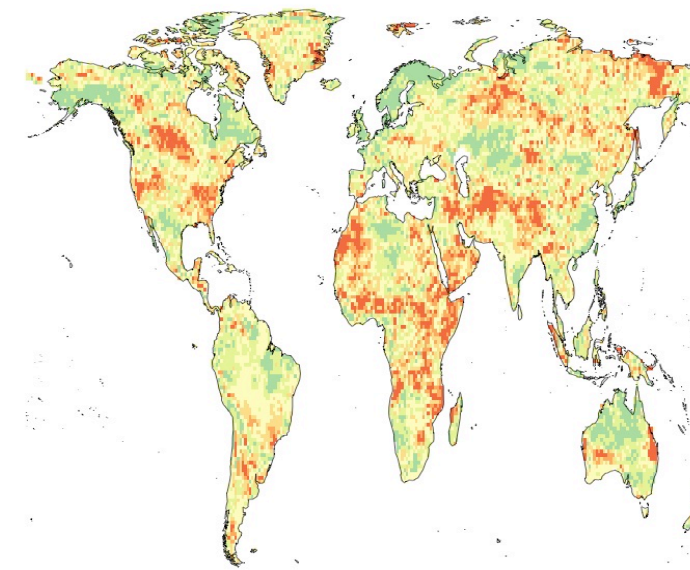
Emulation error



IV of Emulation



IV of Emulation





## From here

Hypothesis/Diagnosis of the reasons for better or worse performance.

Bottom line: Emulation of many metrics of extremes is accurate and can replicate internal variability, thus promises to be an effective and computationally efficient way to include impact-relevant climate information in our integrated models.

More to follow: more metrics, more sectors, more impacts.



# Thank you