Model Diagnostics Task Force

Underestimated runoff declines in Earth System Models (ESMs)

Hanjun Kim, Flavio Lehner (Cornell University)

Andy wood, David Lawrence, Sean Swenson, Katie Dagon, Samar Minallah (NCAR)

2025-02-24

Runoff projections in Earth System Models (ESMs)

- Why do we use ESMs?
  - ESMs simulate the **interactions between climate system**, which is essential for projecting climate change

- For the future water resource assessment, runoff projections from ESMs are being utilized

LETTERS
PUBLISHED ONLINE: 23 DECEMBER 2012 | DOI: 10.1038/NCLIMATE1787
nature climate change

Projections of declining surface-water availability for the southwestern United States

Richard Seager[1]*, Mingfang Ting[1], Cuihua Li[1], Naomi Naik[1], Ben Cook[2], Jennifer Nakamura[1] and Haibo Liu[1]

nature COMMUNICATIONS
https://doi.org/10.1038/s41467-021-25026-3 OPEN

Future global urban water scarcity and potential solutions

Chunyang He[1,2], Zhifeng Liu[1,2]✉, Jianguo Wu[1,2,3], Xinhao Pan[1,2], Zihang Fang[1,2], Jingwei Li[4] & Brett A. Bryan[5]

The global urban population facing water scarcity (CMIP6 runoff):
0.93 billion (2016) → **1.70–2.37** billion people (2050)

Source of uncertainty in regional runoff projections (ΔQ)

Meteorological forcings

Precipitation response (ΔP)

Main driver of runoff
Highly uncertain

Temperature responses (ΔT)

Incomplete proxy for ET
Relatively robust

ΔQ are generally more uncertain than either ΔP & ΔT

Source of uncertainty in regional runoff projections (ΔQ)

Meteorological forcings

Precipitation response (ΔP)

Main driver of runoff
Highly uncertain

Temperature responses (ΔT)

Incomplete proxy for ET
Relatively robust

Additional uncertainties

Sensitivity of runoff to ΔP & ΔT (runoff sensitivity)

The runoff generation process related to warming are complex

Vegetation feedback

Stomatal closure (Q↑)
Betts et al 2007

Vegetation greening (Q↓)
Mankin et al. 2019

snow/glacier melt

Direct increase (Q↑)
Cui et al., 2023

Snow-albedo feedback (Q↓)
Milly and Dunne, 2020

Changes in P Characteristics

Phase shift from snow to rain (Q ↕)
Berghuijs et al., 2014

Increase in extreme p (Q↑)
Wainwright & Parsons, 2002

Changes in seasonality (Q↕)
Scheff et al., 2022

Meteorological forcings

Precipitation response (ΔP)	Temperature responses (ΔT)

Additional uncertainties

Sensitivity of runoff to ΔP & ΔT (**runoff sensitivity**)

Research Objectives

1. Quantify the model bias in runoff sensitivity
2. Using the runoff sensitivity bias, constrain future runoff projections
   → Does the biases matter for the future projection?

Estimation of runoff sensitivity using multiple linear regression

Runoff sensitivity

$$\delta Q \approx \alpha\delta P + \beta\delta T + c\delta P\delta T$$

δ: 5-yr averaged temporal variations

$$\alpha = \frac{\partial(\delta Q)}{\partial(\delta P)} \qquad \beta = \frac{\partial(\delta Q)}{\partial(\delta T)}$$

P sensitivity — T sensitivity

Regression slope is trained for historical period (1948-2017)

P sensitivity: Q changes [%] per unit P increase [%]

T sensitivity: Q changes [%] per unit T increase [K]

T sensitivity is uncertain among climate models

Runoff sensitivity

$$\delta Q \approx \alpha \delta P + \beta \delta T + c\delta P \delta T$$

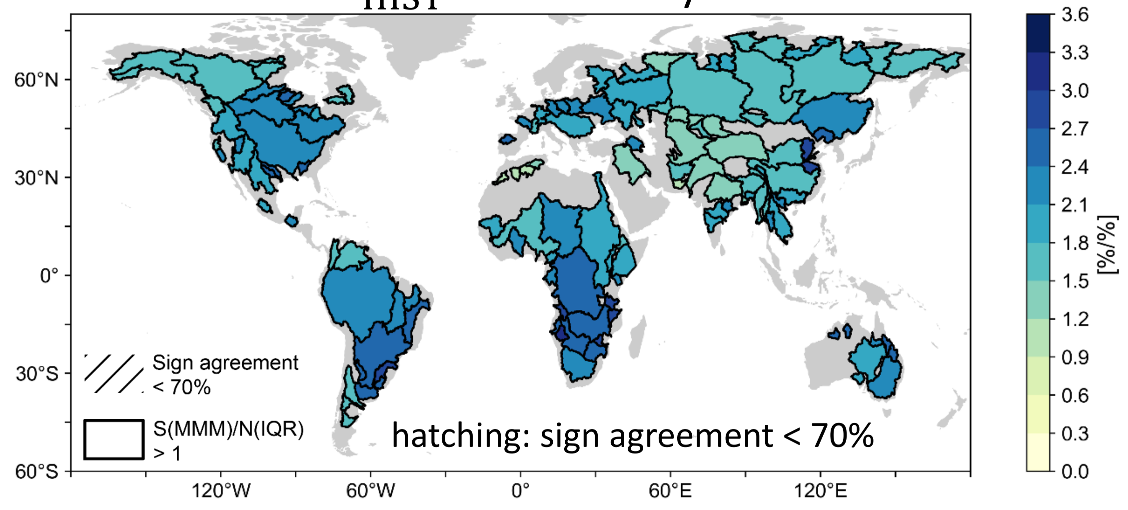
δ: 5-yr averaged temporal variations

$$\alpha = \frac{\partial(\delta Q)}{\partial(\delta P)} \qquad \beta = \frac{\partial(\delta Q)}{\partial(\delta T)}$$

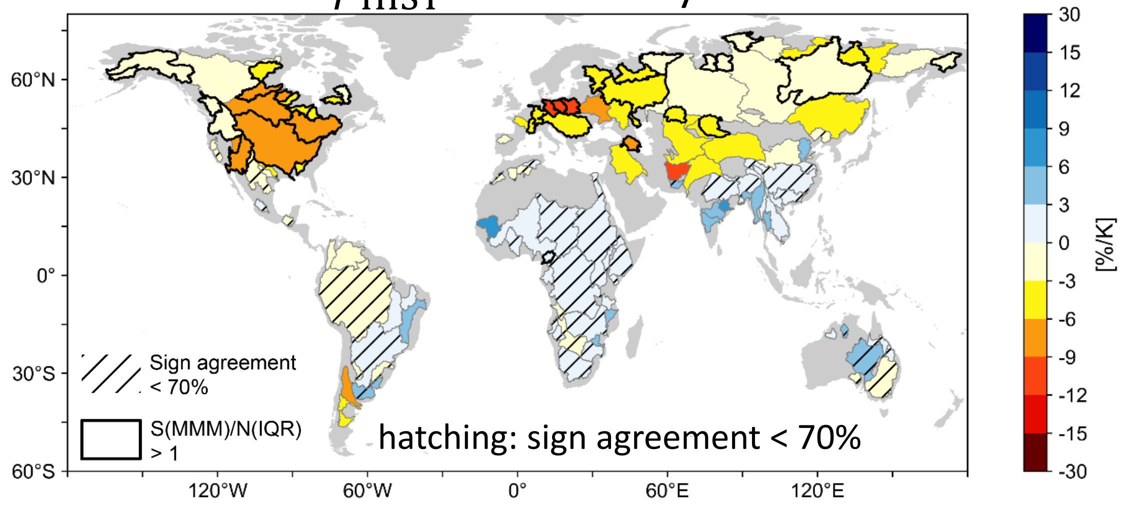
P sensitivity T sensitivity

MMM runoff sensitivity (28 CMIP6 models, 1948-2017)

α_{HIST}: P sensitivity



β_{HIST}: T sensitivity



Does this runoff sensitivity capture the effects of climate change on runoff generation?

Runoff sensitivity from historical simulation

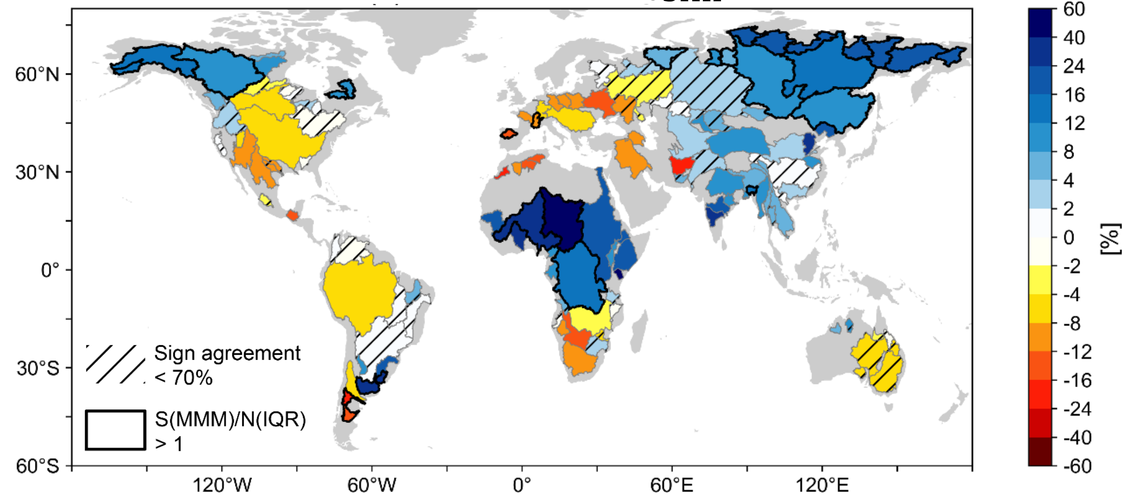
$$\delta Q = \alpha_{HIST}\delta P + \beta_{HIST}\delta T$$

Prediction of each model's runoff projection

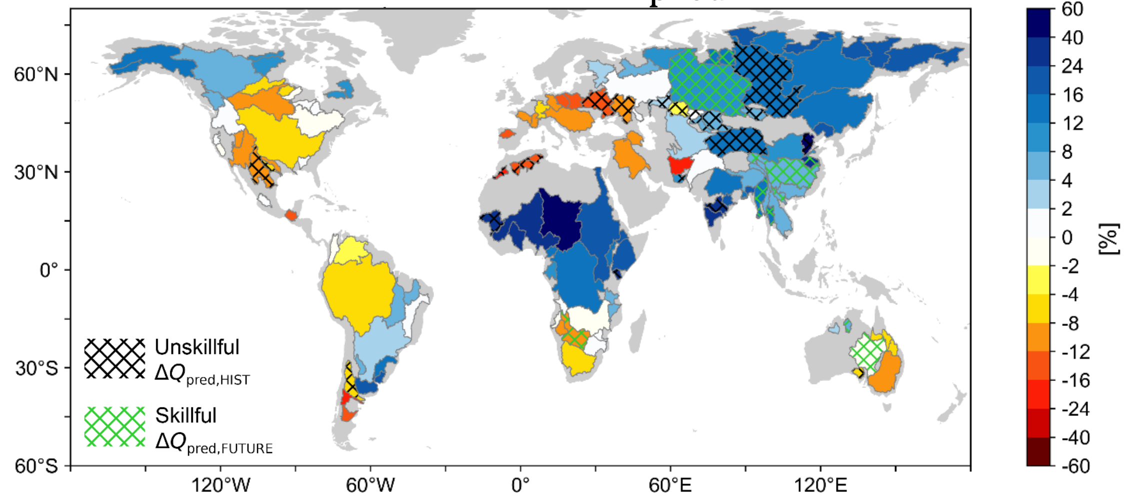
$$\Delta Q_{pred} = \alpha_{HIST}\Delta P + \beta_{HIST}\Delta T$$

Δ: SSP245 future changes (2030-2070)

CMIP6 MMM ΔQ_{sim}



CMIP6 MMM ΔQ_{pred}



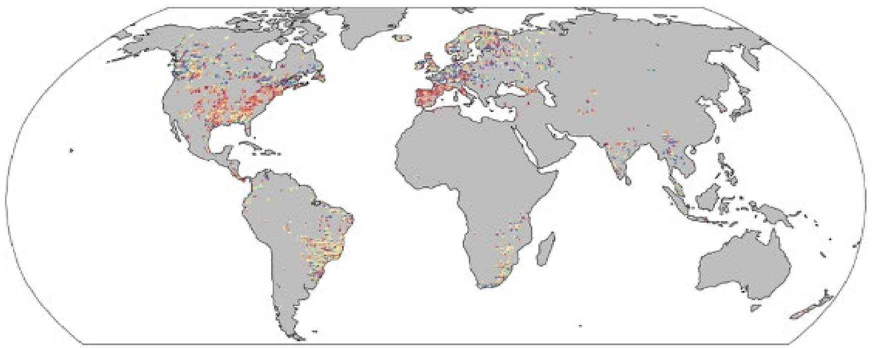
✓ Statistically indistinguishable multi-model median + Significant inter-model correlation

→ The historical runoff sensitivity can skillfully predict the future changes in 97 among 131 basins.

How biased is the model sensitivity?

- GRUN as OBS proxy: ML-based global runoff reanalysis dataset

GSIM-derived training data (5094 grid cell)



G-RUN ENSEMBLE: A Multi-Forcing Observation-Based Global Runoff Reanalysis 1903-2017 monthly / 0.5° × 0.5°

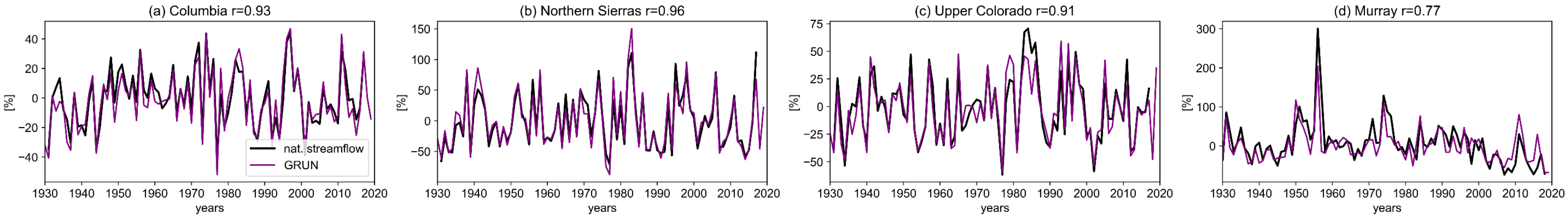
G. Ghiggi[1,2], V. Humphrey[3,4], S. I. Seneviratne[2], and L. Gudmundsson[2]

$$Q = \text{Random Forest Model}\,(P_{\text{past 6 months}},\ T_{\text{past 6 months}})$$

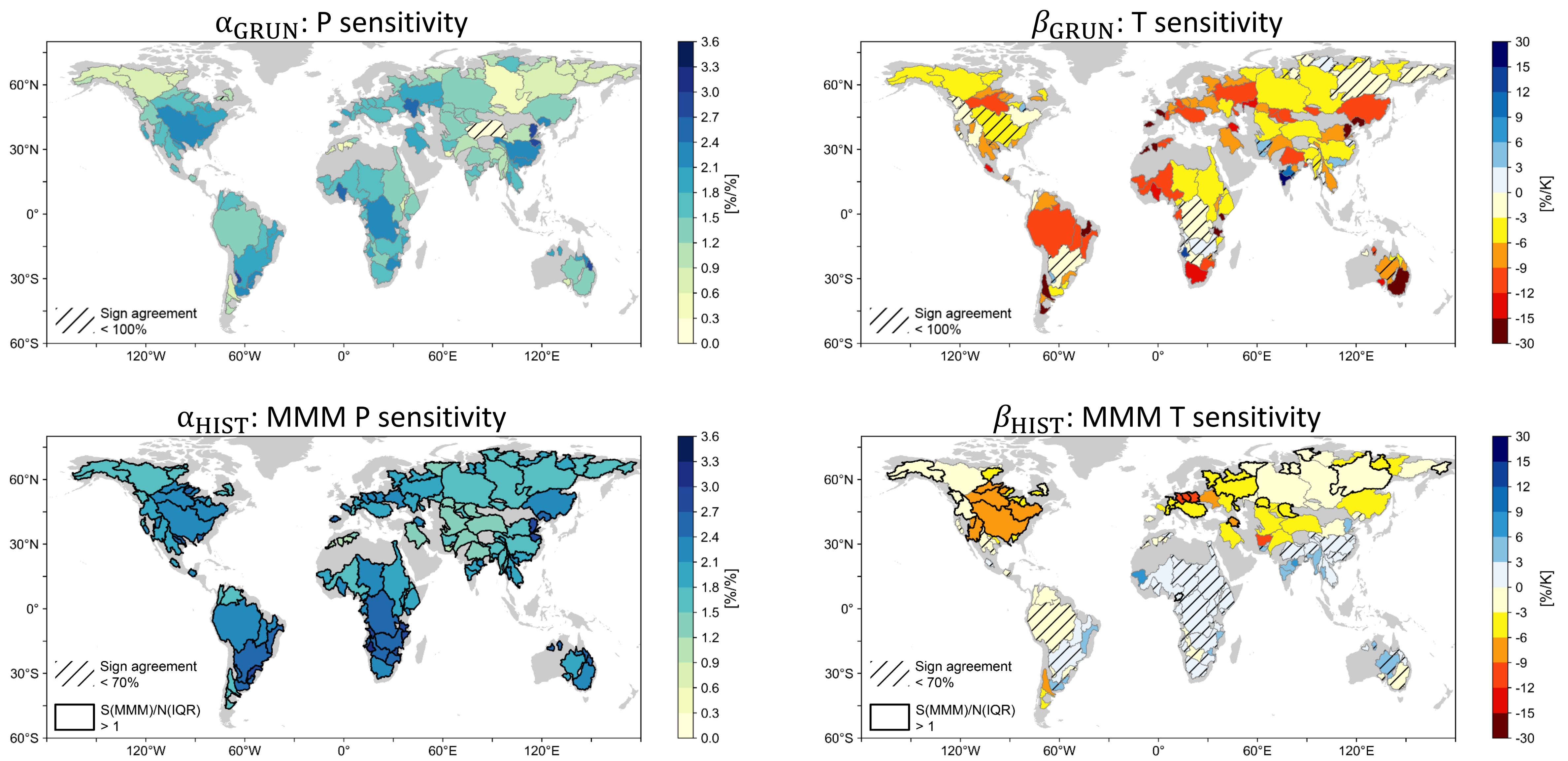
✓ 100 ensembles members exist, enabling the quantification of observational uncertainty

✓ GRUN data is shown to be outperforming the reanalysis data; However, the data is still not reliable for some basins.

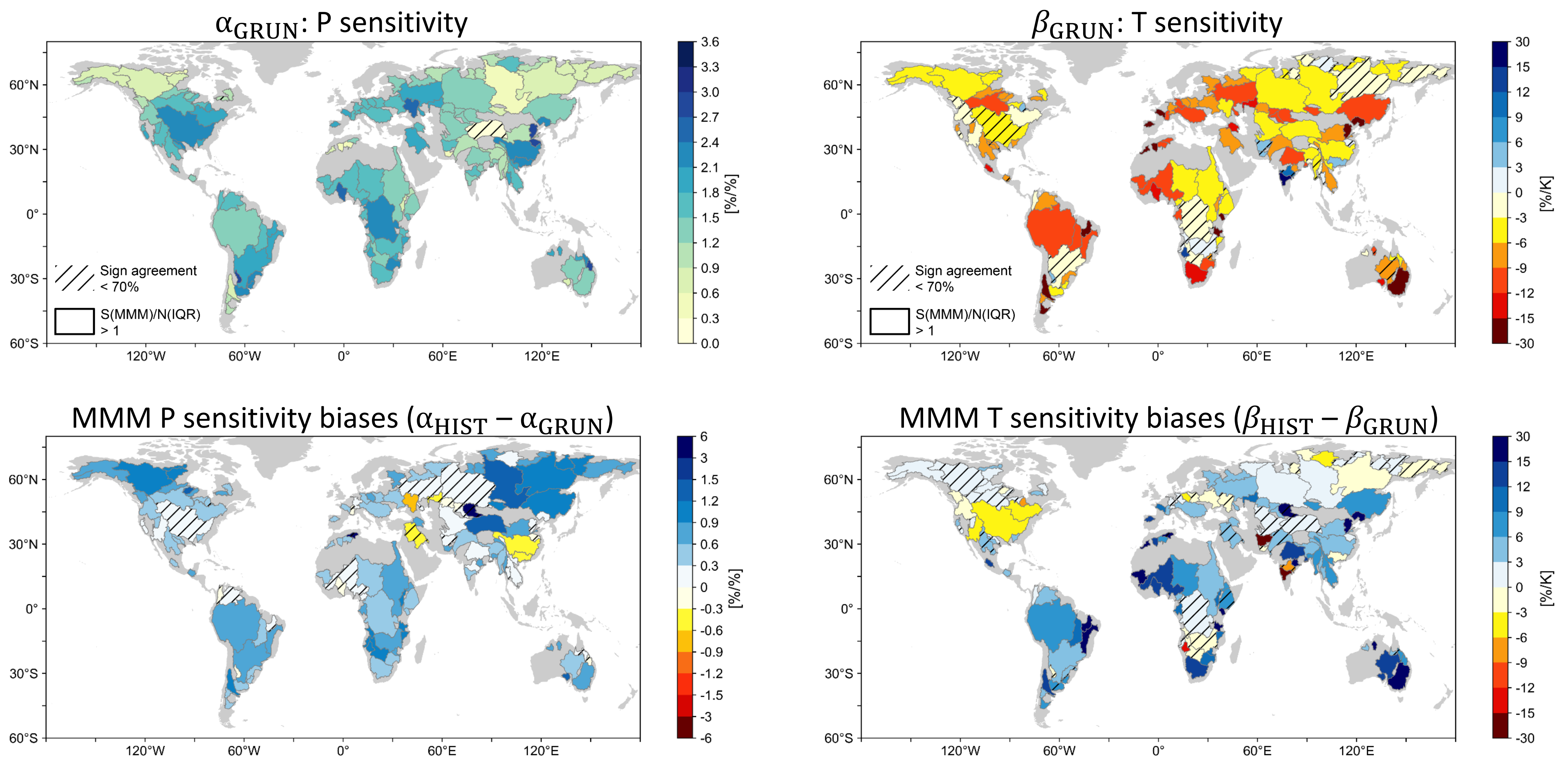
Annual timeseries of runoff



How biased is the model sensitivity?



In models, P sensitivity is more positive, and T sensitivity is less negative



How much these biases affect the future projections? → Observational constraint

Prediction using runoff sensitivity

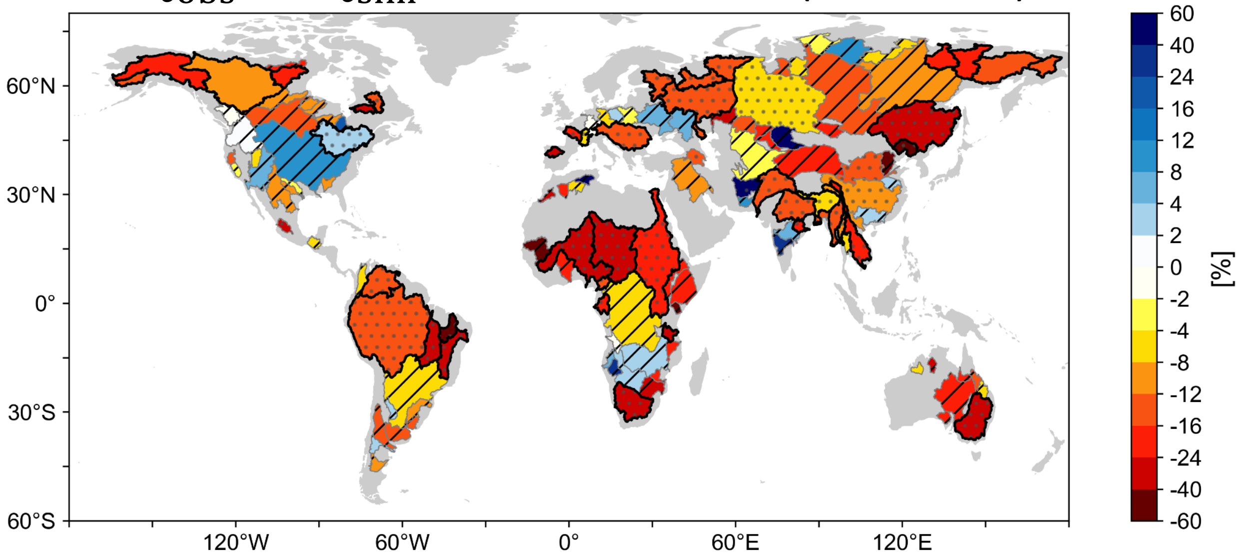
$$\Delta Q_{pred} = \alpha_{HIST}\Delta P + \beta_{HIST}\Delta T$$

Δ: SSP245 future changes (2030-2070)

Observationally-constrained projection

$$\Delta Q_{obs} = \alpha_{GRUN}\Delta P + \beta_{GRUN}\Delta T$$

$\Delta Q_{obs} - \Delta Q_{sim}$: SSP245 MMM OC (28 models)



✓ Overall, the observationally-constrained projections indicate a drier future than the unconstrained projections

✓ The correction effect mainly arises from the T sensitivity bias

How much these biases affect the future projections? → Observational constraint

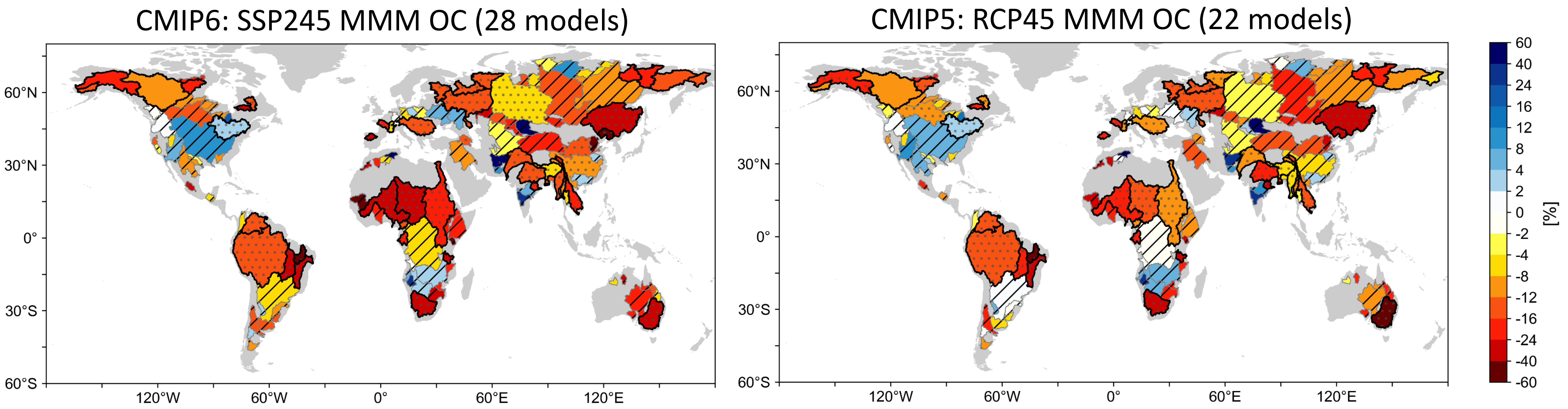
Prediction using runoff sensitivity

$$\Delta Q_{pred} = \alpha_{HIST}\Delta P + \beta_{HIST}\Delta T$$

Δ: SSP245 future changes (2030-2070)

Observationally-constrained projection

$$\Delta Q_{obs} = \alpha_{GRUN}\Delta P + \beta_{GRUN}\Delta T$$



✓ Inter-basin correlation = 0.91 → the constraining effect is consistent regardless of model generation

The observational constraint for central value is robust for 41 of 131 global river basins

Is the correction effect statistically significant?

$\Delta Q_{pred} == \Delta Q_{sim}$
Skillful prediction

(107/131 basins)

$\Delta Q_{obs} \neq \Delta Q_{sim}$
Discernible correction

(68/107 basins)

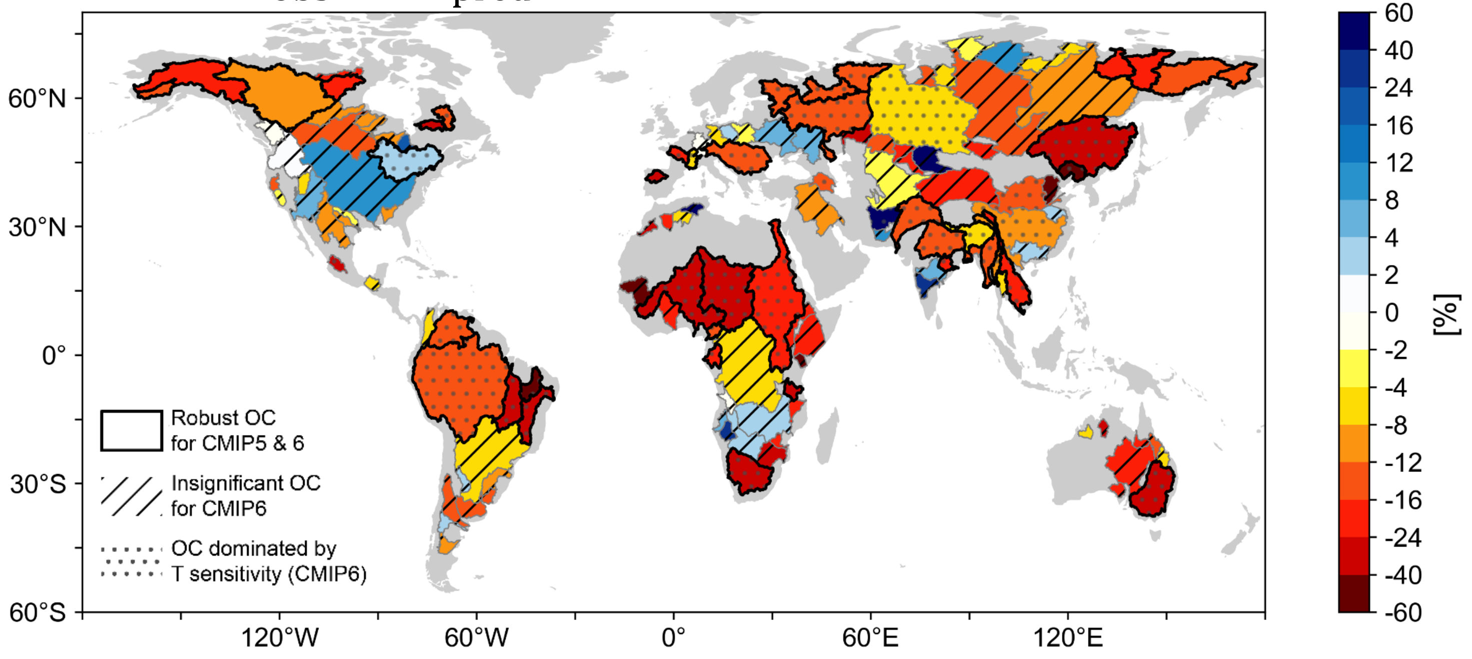
$\Delta Q_{obs} - \Delta Q_{sim} > NS + IV$
Smaller non-stationarity and internal variability

(51/68 basins)

Significant corrections both in CMIP5 and CMIP6
→ 'robust constraint'

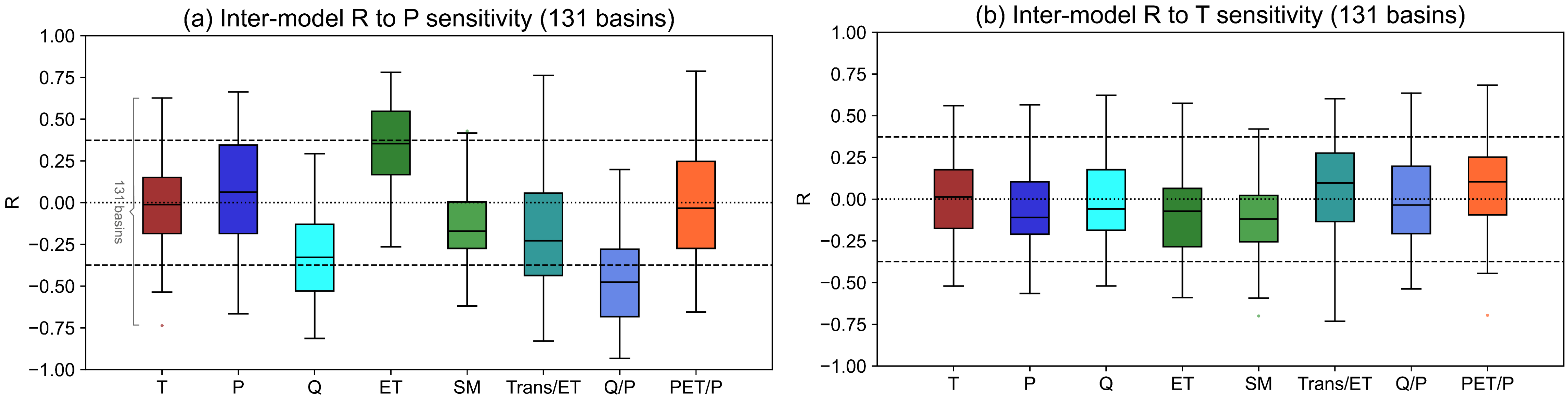
(41/51 basins)

$\Delta Q_{obs} - \Delta Q_{pred}$: SSP245 MMM OC (28 models)



What is the cause of the sensitivity bias?

For each basin, we get inter-model correlation between mean state variables and runoff sensitivity



- ✓ Fixing the bias in runoff ratio (Q/P) may improve P sensitivity bias.
- ✓ It is unlikely to resolve the more critical T sensitivity bias if we only focus on mean state bias.

Key point

"The runoff decline due to temperature increase is generally underestimated in ESMs"

This bias can be quantified by the T sensitivity using historical timeseries

Discussions

1. The degree of sensitivity bias is affected by observational dataset, but overall results are consistent

When validated with more reliable station data for selective basins, T sensitivity biases are more negative.
The underestimated drying in ESMs are also consistent to other studies using different datasets or statistical methods

Zhang et al. 2023, Douville 2024

2. To help reducing the T sensitivity bias, we have implemented runoff sensitivity metrics to diagnostic package (NOAA MDTF & ILAMB)

As traditional modeling approach focusing on mean state biases would not resolve T sensitivity bias,
we added the model diagnostics of runoff sensitivity to the NOAA MDTF metrics package and ILAMB package

Example: NOAA MDTF diagnostic package (CESM2)

runoff sensitivity diagnostic

C:/Users/reapr/내%20드라이브/Research/runoff_project/MDTF/ver241202/runoff_sensitivities/runoff_sensitivities.html

NOAA
Model Diagnostics Task Force (MDTF)
Diagnostics Package

Runoff sensitivity: The sensitivity of runoff to surface air temperature and precipitation

General description

The sensitivity of runoff (Q) to temperature (T sensitivity; %/K) and precipitation (P sensitivity; %/%) are calculated using the multiple linear regression, using the 5-year averaged annual timeseries of the historical period during 1945-2014. The sensitivity is calculated respectively for 131 global river basins, after averaging the variables for water year (Oct. to Sep.) and each basin using the GRDC Major River Basin masks

The calculated runoff sensitivity is then compared with the pre-calculated sensitivities for observations (GRUN-ENSEMBLE), CMIP5/6 historical simulations. GRUN-ENSEMBLE is machine-learning based global runoff reanalysis, providing 100 realizations to sample to observational uncertainty (Ghiggi et al. 2021). For CMIP models, we use historical simulation from 1945 to 2014. For CMIP5, we extend the timeseries to 2014 using RCP4.5 scenario. The ranges from observational uncertainty, inter-model spread, and uncertainty of regression coefficients are shown for individual river basins.

In CMIP5 and CMIP6, the uncertainty of T sensitivity can contribute to the runoff projection uncertainty as much as the different P projections among models. The runoff sensitivity in climate model is often biased; the negative T sensitivity is often too weak in climate models, indicating the underestimation of future runoff declines.

While the P sensitivity is generally correlated with the mean state biases, the T sensitivity exhibits no systematic relationship with mean state variables. Hence, the traditional modeling approaches, which focus on improving mean state biases, may not reseolve the biases in T sensitivity. Therefore, the new diagnostic metric related to the runoff sensitivity is needed for future land model development.

References

Calculation and interpretation of runoff sensitivities:

Lehner, F., Wood, A. W., Vano, J. A., Lawrence, D. M., Clark, M. P., and Mankin, J. S. (2019). The potential to reduce uncertainty in regional runoff projections from climate models. Nature Climate Change, 9(12), 926-933.
Kim, H., Lehner, F., Dagon K., Lawrence, D. M., Swenson, S., Wood, A. W. (2025). Constraining climate model projections with observations amplifies future runoff declines. in preparation.

Observational estimate (GRUN-ENSEMBLE):

Ghiggi, G., Humphrey, V., Seneviratne, S. I., & Gudmundsson, L. (2021). G-RUN ENSEMBLE: A multi-forcing observation-based global runoff reanalysis. Water Resources Research, 57(5), e2020WR028787.

Caution

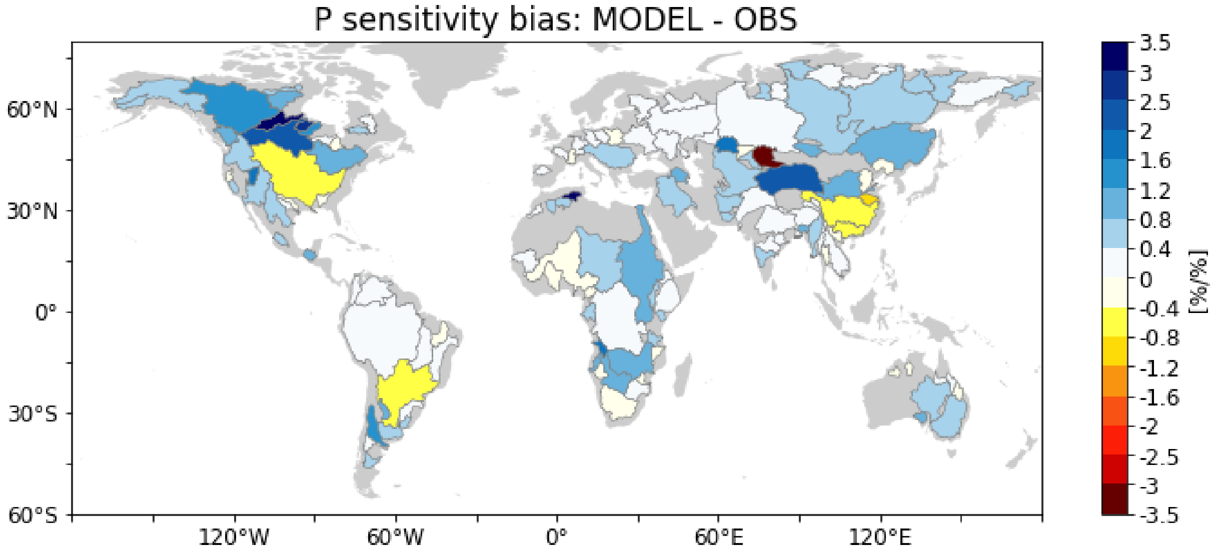
For some models, the surface water budget (Precipitation - Evaporation = Runoff) is not closed in the long-term average.
To evaluate the water budget closure (WBC) for the period of {{FIRSTYR}}-{{LASTYR}}, we specify river basins in which WBC is kept within 10% error. The fraction of closed river basins compraed to all river basins is used as criteria.
For example, WBC = 0.91 for CESM2, and it means that WBC is kept for 91% of the area of global river basins.
if WBC < 0.6, we consider the model's runoff data to be inappropriate for runoff sensitivity calculations. We do not know why the water balance is not closeable in some models; it might have more to do with the data they provide in the CMIP archive rather than the model itself.
WBC value is shown in the plots of MODEL runoff sensitivities (title).

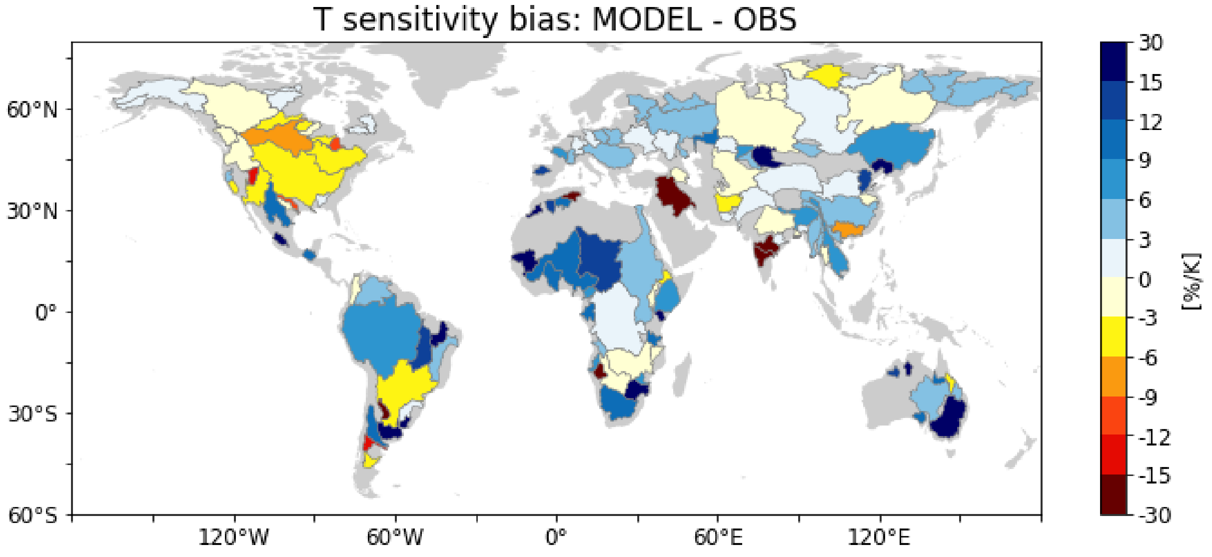
Casename: {{CASENAME}}

Runoff sensitivities	OBS	MODEL	MODEL-OBS	CMIP6-OBS
T sensitivity (%/K)	plot	plot	plot	plot
P sensitivity (%/%)	plot	plot	plot	plot

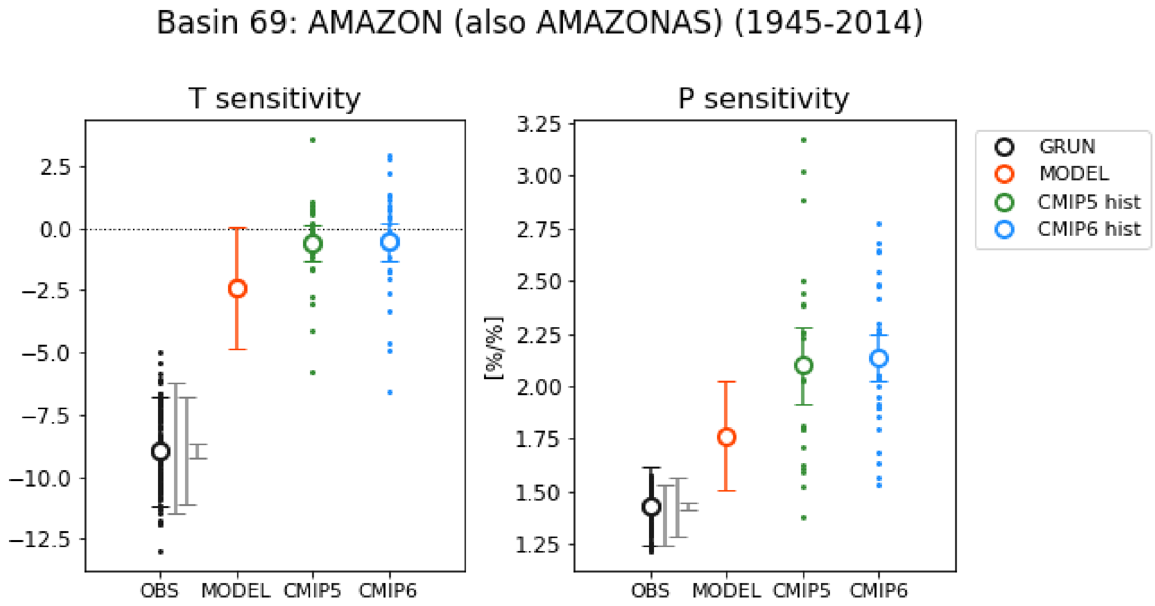
Summary

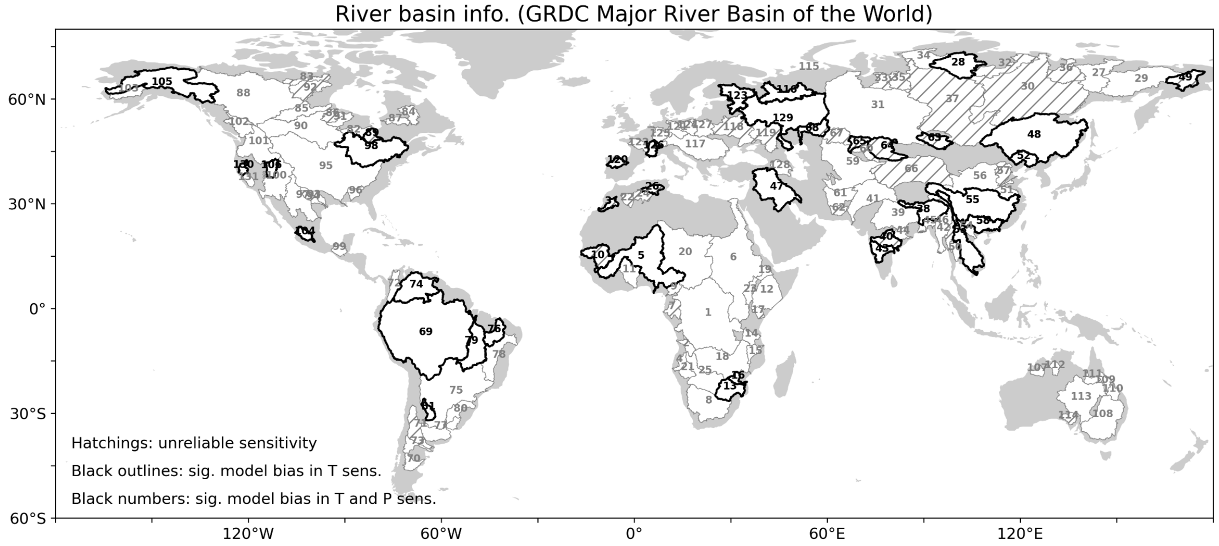
NOAA MDTF diagnostic package output for CESM2





NOAA MDTF diagnostic package output for CESM2





Key point

"The runoff decline due to temperature increase is generally underestimated in ESMs"

This bias can be quantified by the T sensitivity using historical timeseries

Thank you!

Q&A

Contact: hk764@cornell.edu