



**University
of Victoria**



**PACIFIC CLIMATE
IMPACTS CONSORTIUM**

Quantifying the human influence on the intensity of extreme 1- and 5-day precipitation amounts at global, continental, and regional scales

Qiaohong Sun¹, Francis Zwiers^{1, 2}, Xuebin Zhang³, Jun Yan⁴

¹ Pacific Climate Impacts Consortium, University of Victoria, Victoria, British Columbia V8W 2Y2, Canada.

² Nanjing University of Information Science and Technology, Nanjing, China;

³ Climate Research Division, Environment and Climate Change Canada, Toronto, Ontario M3H 5T4, Canada

⁴ Department of Statistics, University of Connecticut, Storrs, CT



1 Introduction

- ❖ Previous detection and attribution analyses suggest that human-induced increases in greenhouse gases have contributed to observed changes in extreme precipitation.
- ❖ However, all previous detection and attribution studies of observed changes in extreme precipitation:
 - (1) use station data that have been heavily processed via gridding, transformation, and spatial and temporal averaging or other dimension reduction approaches, as well as using climate models to estimate the responses to external forcing,
 - (2) use models to estimate the unforced natural variability of extreme precipitation.
- ❖ Both aspects reduce user confidence in detection and attribution results.



Research question

It is important to ask:

- if human influence on extreme precipitation at global and continental scales is robustly detectable using different datasets, data processing methods, different detection and attribution methods, and different models.
- if external influence can be detected at the regional scale?

Two extensions:

- Propose a novel detection and attribution analysis method that is applied directly to station data in the areas considered without prior processing and use climate models only to obtain estimates of the space-time pattern of extreme precipitation response to external forcing. Changes at stations are evaluated by fitting non-stationary Generalized Extreme Value (GEV) distributions at individual stations to the observed values.
- Perform detection and attribution analysis on extreme precipitation at different spatial scales including global, continental, and regional scales.

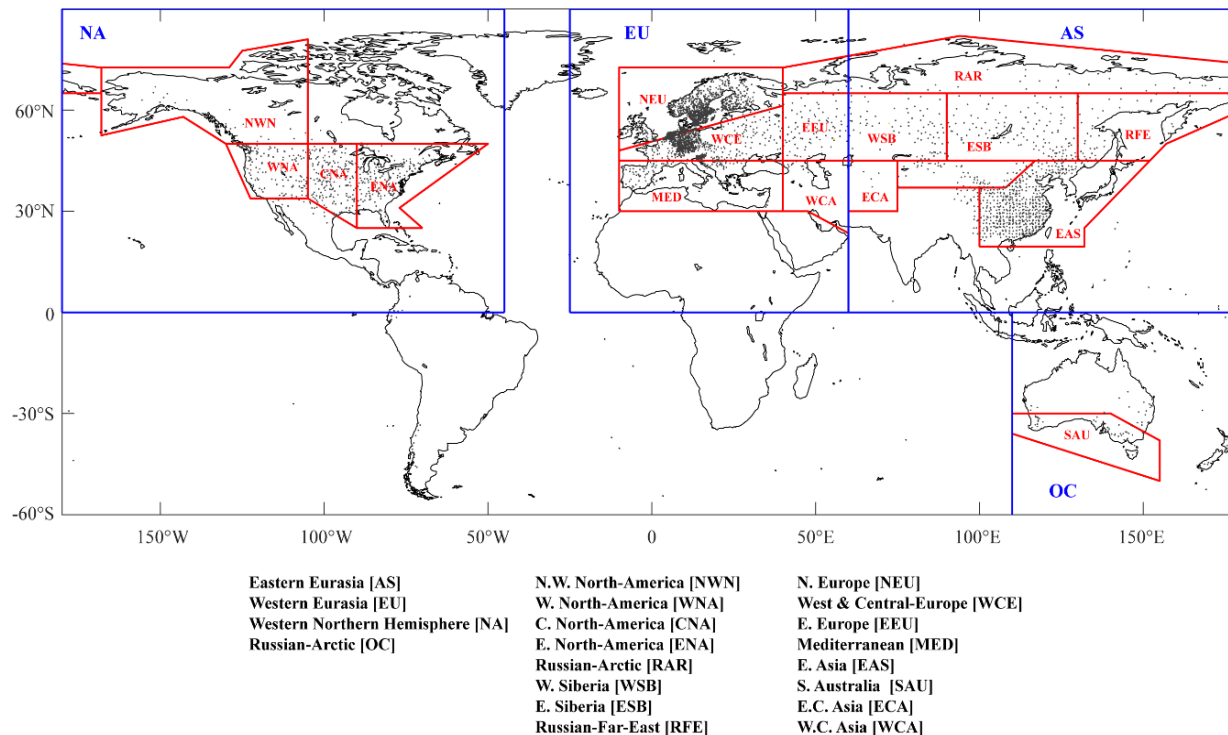


2 Data

➤ Observations

Rx1day and Rx5day from a stations-based observational dataset, which includes 14,796 land-based observing stations with more than 30 years of record during 1900 to 2018 across the globe (Sun et al. 2021).

To consider long-term changes, we selected 5,081 stations with more than 45 years of data during the period 1950-2014, including at least 3 years during 2010-2014.





2 Data

➤ Model simulations

CanESM2 Large ensemble

50 perturbed initial condition members - one with anthropogenic and natural forcings combined (ALL forcings) from 1950 to 2020

CMIP6 multi-model ensemble

Simulations from ten CMIP6 models (Table 1) that provide data from at least three ensemble members for both ALL and NAT forcing.

Model	ALL	NAT	Ensemble Size
ACCESS-CM2	✓	✓	3
BCC-CSM2-MR	✓	✓	3
CanESM5	✓	✓	10
CNRM-CM6-1	✓	✓	3
CESM2	✓	✓	3
HadGEM3-GC31	✓	✓	4
IPSL-CM6A-LR	✓	✓	9
MIROC6	✓	✓	3
MRI-ESM2-0	✓	✓	5
NorESM2-LM	✓	✓	3
Total	10 models, 46 runs for each forcing		



3 Method

➤ Detection and attribution method

In order to take advantage of the possibility that both observed and modelled precipitation extremes may scale similarly with temperature, we considered fitting GEV distributions to log transformed observations, $\ln(\text{Rx1day})$ and $\ln(\text{Rx5day})$, with GEV location parameters of the form

$$\mu_{ts} = \mu_{bs} + \sum_{i=1,\dots,l} \beta_i \ln(\gamma_{tsi}) \quad (1)$$

Where, $\gamma_{ts} = (\gamma_{ts1}, \dots, \gamma_{tsl})^T$ is an $l \times 1$ vector containing the signals from l forcings of interest from climate model output, which can be expressed as the ratio between the median of Rx1day or Rx5day at time t and that at a fixed reference time b .

We assume that the scale and shape parameters σ_s and ξ_s vary only with geographical location s .

Here, the noise in the γ_{ts} estimates could be reduced by averaging over large regions since we could expect homogeneous scaling over very broad regions:

- Scaling research and model assessment papers (e.g. Kharin et al. 2013, 10.1007/S10584-013-0705-8; Li et al. 2020, 10.1175/JCLI-D-19-1013.1) suggests that γ_{ts} varies smoothly in space, at least for the ALL, anthropogenic (ANT), and NAT signals.
- There are relatively little data in areas where these factors are expected to increase relatively quickly over time well past one (such as tropical convergence zones) or to decrease over time (such as regions affected by changes in the downwelling branches of the Hadley circulation).



3.1. Signal estimation from ensemble simulations

A $k \times k$ grid box moving window is used to smooth the signal across space and a quadratic b-spline is used to smooth the signal across time.

For the CanESM2 large ensemble, raw signal estimation at each location from an individual climate model using a specified forcing (described for Rx1day) is as follows:

- For the entire base period 1950-2014, pool Rx1day values from within a $k \times k$ grid box region from all ensemble members and estimate the median;
- For each year, pool Rx1day values from the $k \times k$ grid boxes and from all ensemble members; estimate the median and divide by the median for the base period.
- Apply a quadratic b-spline with knots placed every N years to the raw signal estimates to obtain the smoothed signals $\hat{\gamma}_{ts}$ across time.

For the multi-model CMIP6 ensemble

- For each year, interpolate the raw signals from individual models to a common grid;
- Average the model signals to obtain a multi-model signal field on the interpolation grid;
- At each location on the interpolation grid, fit a B-spline with knots at every N years for that location to obtain smoothed signal estimates $\hat{\gamma}_{ts}$.

Here, we use $k=3$, place spline knots at 5-year intervals



3.2. Detection analysis

A **combined score equation (CSE) method** is used to estimate the scaling factors in model (1). Specifically, at time t and location s , let $f(\cdot; \mu_{ts}, \sigma_s, \xi_s)$ be the density function of the GEV distribution with location, scale and shape parameters μ_{ts} , σ_s , and ξ_s respectively, and let $S_{ts} = d \log f(\ln(Y_{ts}); \mu_{ts}, \sigma_s, \xi_s) / d\mu_{ts}$. Then the CSE for β over a domain with m stations is

$$\sum_{t=1}^n \ln(\gamma_t) W_t^{-1} S_t = 0 \quad (2)$$

where $\gamma_t = (\gamma_{t1}, \dots, \gamma_{tm})$, W_t^{-1} is $m \times m$ weight matrix, and $S_t = (S_{t1}, \dots, S_{tm})^T$. Estimating equations of site-specific parameters $(\mu_{bs}, \sigma_s, \xi_s)$ are based on the derivatives of the log likelihood with respect to these parameters, which only depend on data at each site s .

Here, we opt not to explicitly model spatial dependence and thus use the special case with the identity weight matrix, referred to as the **CSE-I method**, for estimating the scaling factors in model (1).



3.3. Confidence intervals

Confidence intervals are constructed for the scaling factor estimates by using a two-level block bootstrap procedure that accounts for uncertainty in signal and scaling factor estimation.

- ***Signal bootstrap***

For each climate model, resampling is performed to generate additional bootstrap samples of its ensemble of simulations in a way that retains temporal signal structure, spatial dependence, and as much temporal dependence as possible. Each such bootstrap sampled ensemble is then used to produce a bootstrap estimate of that model's smoothed signal.

- ***Scaling factor uncertainty estimation***

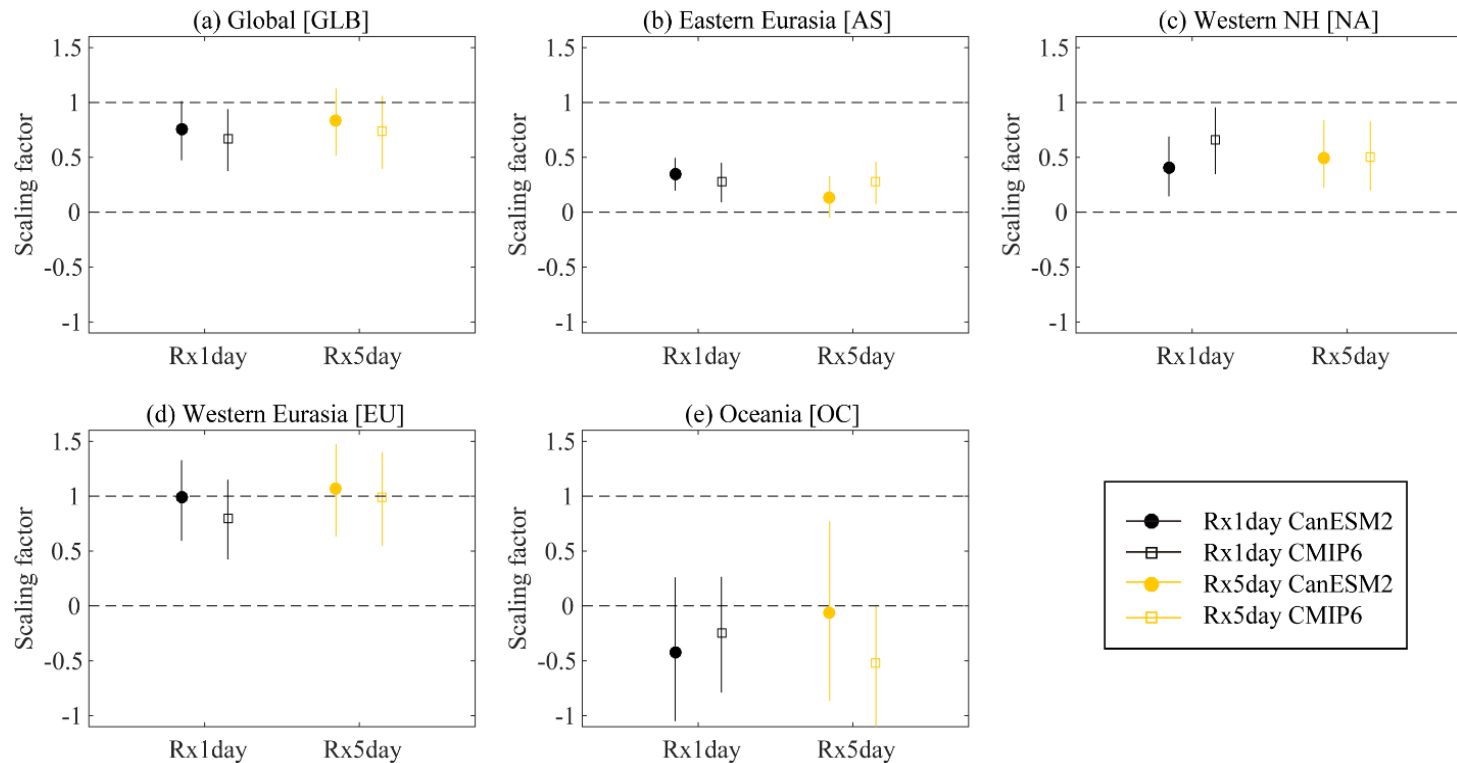
The second level of bootstrapping involves repeatedly fitting the bootstrap estimates of CanESM2 or CMIP6 ensemble signals to the observations of $\ln(Rx1day)$ and $\ln(Rx5day)$ to estimate scaling factors, bootstrapping residuals after removal of the scaled signals, and refitting signals.

4 Results

Global and continental scales



One-single analysis

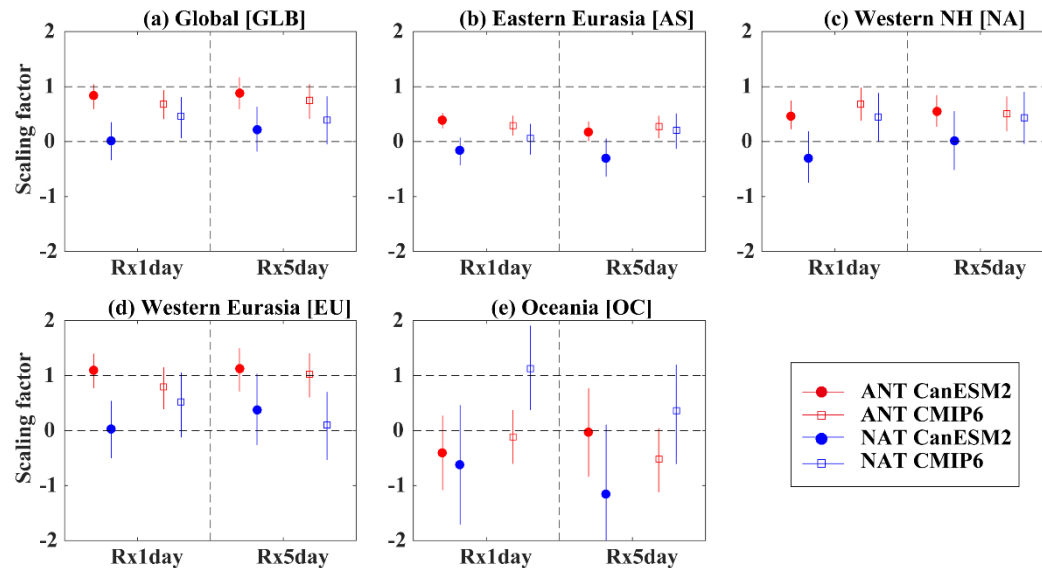


Point estimates of the scaling factors of Rx1day and Rx5day and the corresponding 90% confidence intervals for one-signal (ALL forcings) detection



Two-single analysis

Our results



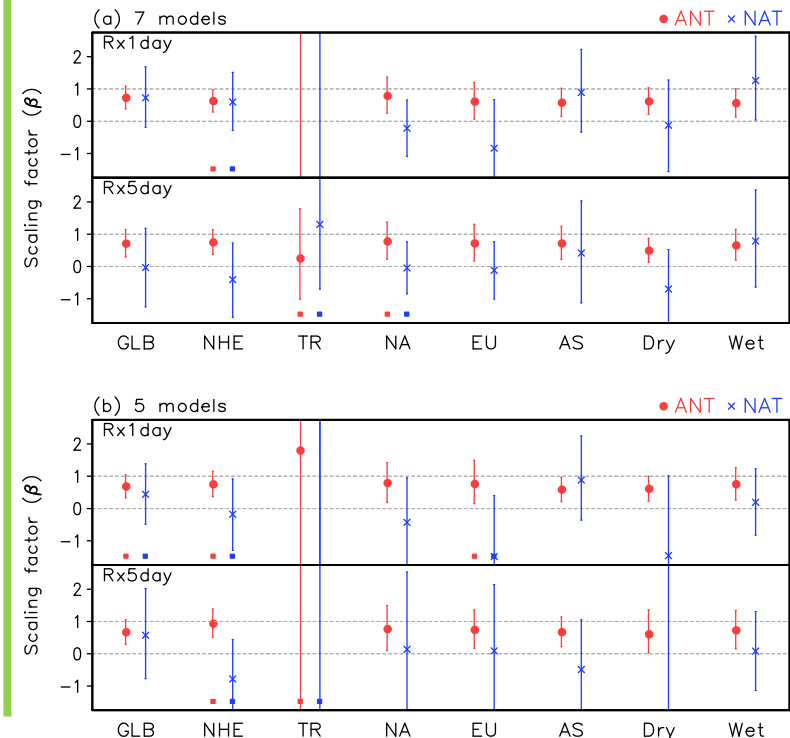
ANT: Anthropogenic forcing

NAT: Natural forcing

Results from Paik et al. 2020

(10.1029/2019GL086875)

Using “probability-based index” (PI)
and Optimal Fingerprinting method

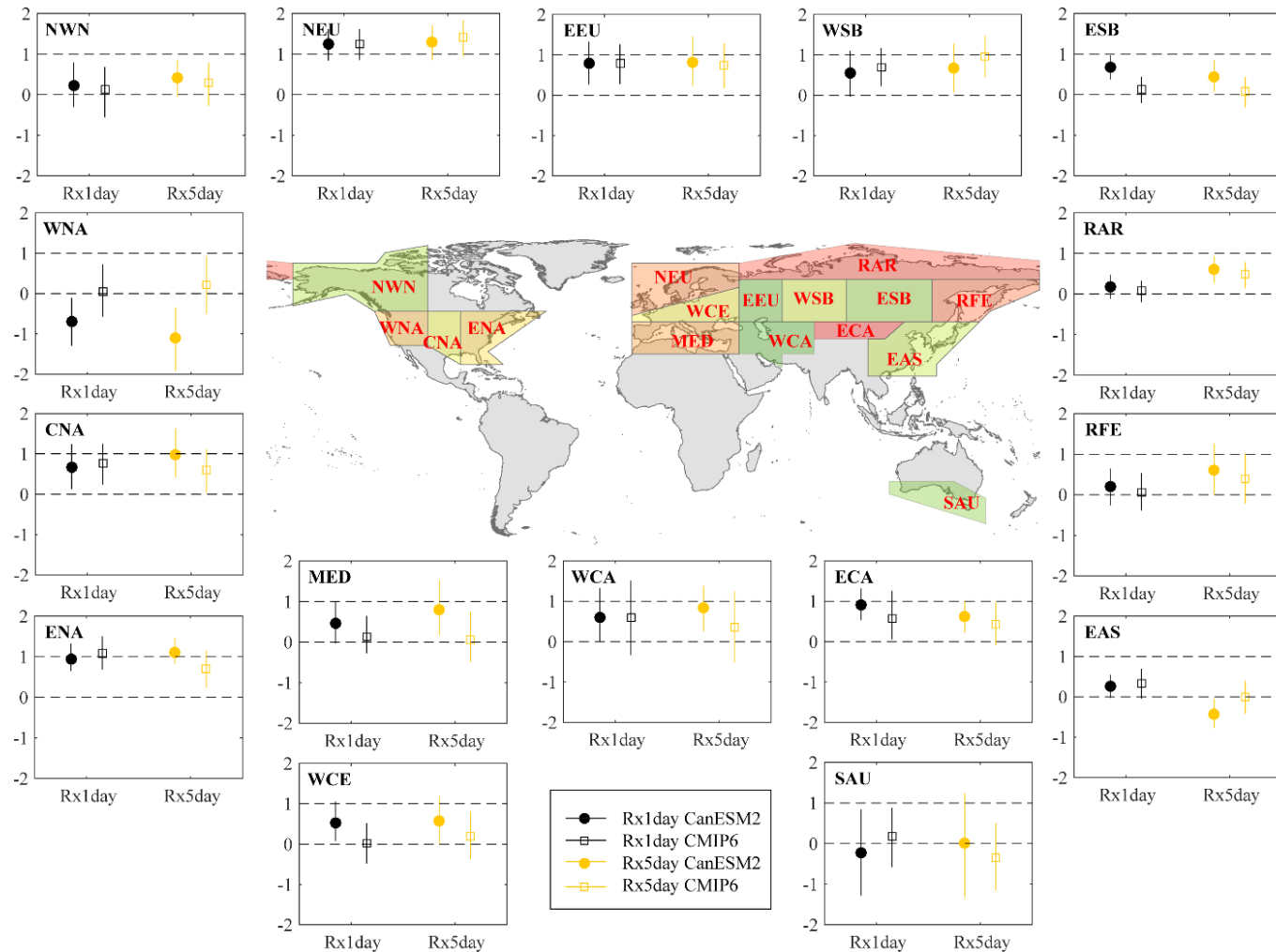


Point estimates of the scaling factors of Rx1day and Rx5day and the corresponding 90% confidence intervals for two-signal (ANT and NAT) detection

Regional scale



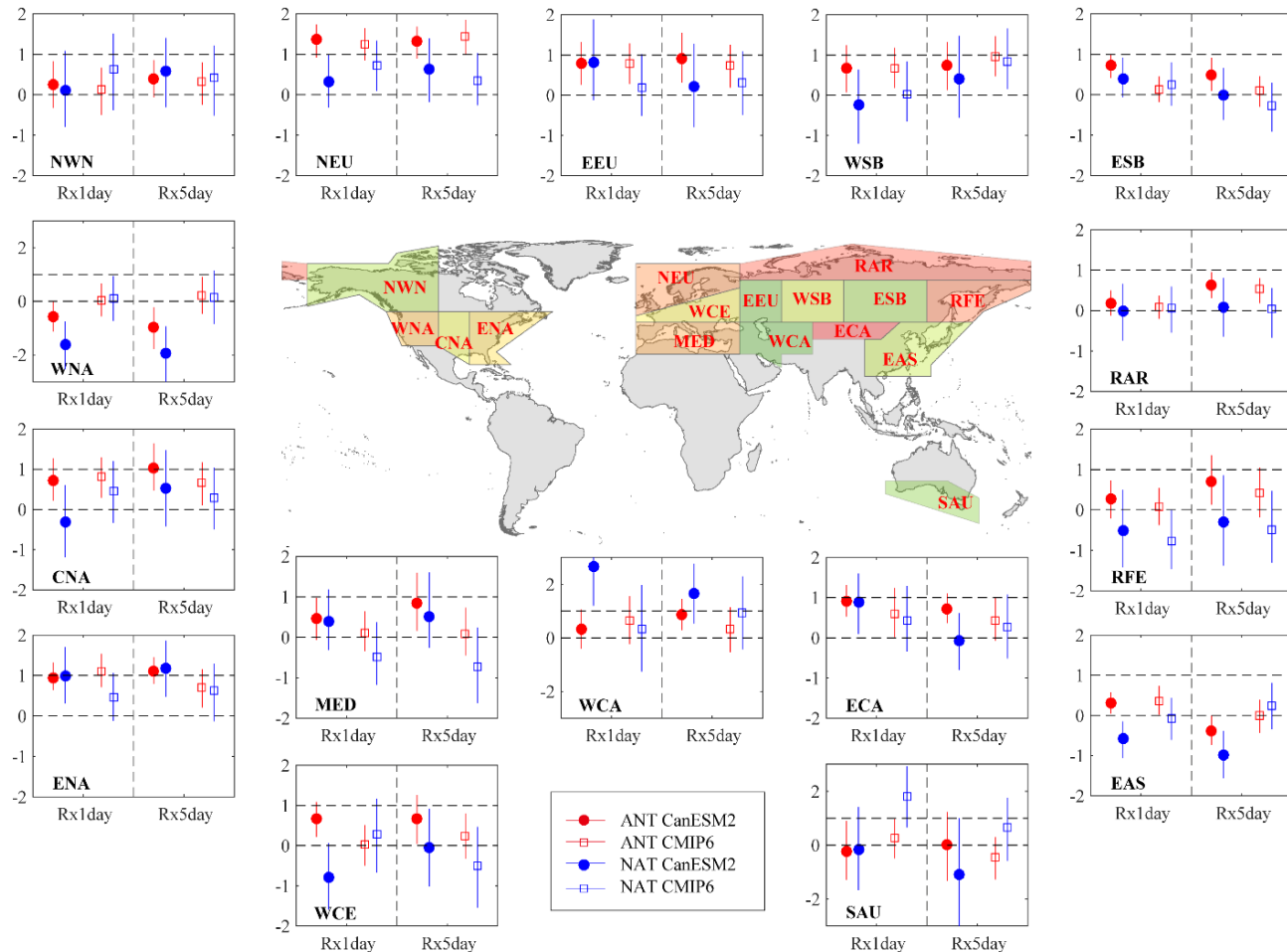
One-single analysis



Point estimates of the scaling factors of Rx1day and Rx5day and the corresponding 90% confidence intervals for one-signal (ALL) detection



Two-single analysis



ANT: Anthropogenic forcing **NAT: Natural forcing**

Point estimates of the scaling factors of Rx1day and Rx5day and the corresponding 90% confidence intervals for two-signal (ANT and NAT) detection



Conclusion

Whether human influence on extreme precipitation at global and continental scales is robustly detectable using different datasets, data processing methods, different detection and attribution methods, and different models?

- Consistency between previous studies (such as Paik et al. 2020) and our study using different data processing methods, different detection and attribution methods, and different models substantially increases confidence in the finding that human influence on extreme precipitation is now robustly detectable in land-based precipitation observations over global land and three continents, including AS, NA and EU.



Conclusion

Whether external influence can be detected at the regional scale?

At the regional scale, external influence can be detected robustly over C. North-America, E. Asia, E.C. Asia, E. Europe, E. North-America, N. Europe, and W. Siberia for Rx1day, and over C. North-America, E. Europe, E. North-America, N. Europe, Russian-Arctic, and W. Siberia for Rx5day.