This form documents the artifacts associated with the article (i.e., the data and code supporting the computational findings) and describes how to reproduce the findings.

Part 1: Data

☐ This paper does not involve analysis of external data (i.e., no data are used or the only data are generated by the authors via simulation in their code).

☒ I certify that the author(s) of the manuscript have legitimate access to and permission to use the data used in this manuscript.

Abstract

The dataset consists of observed daily precipitation data in millimeters from 125 monitoring stations in the Danube river basin (Europe) and from 2229 monitoring stations in the Mississippi river basin (North America) over the period from 1965 to 2020. The temperature covariate used to fit the model in this project was derived from the ERA5-Land reanalysis data for the corresponding region we selected, which is a global land-surface dataset with quite a high spatial resolution of 9km, and the projected temperature covariate is derived from climate models outputs of the sixth Coupled Model Intercomparison Project (CMIP6), namely AWI, MIROC, and MPI.

Availability

☒ Data **are** publicly available.
☐ Data **cannot be made** publicly available.

If the data are publicly available, see the *Publicly available data* section. Otherwise, see the *Non-publicly available data* section, below.

Publicly available data

☒ Data are available online at:

- Climate model outputs and EAR5-Land Reanalysis data: https://www.copernicus.eu/en/access-data
- Precipitation observations: https://www.ncei.noaa.gov/products/land-based-station/global-historical-climatology-network-daily
- Watershed Boundary Dataset: https://www.usgs.gov/national-hydrography/watershed-boundary-dataset
- The data files for plots are stored on Github: https://github.com/PangChung/ExtremePrecip

☐ Data are available as part of the paper's supplementary material.

☐ Data are publicly available by request, following the process described here:

☒ Data are or will be made available through some other mechanism, described here:

- We have saved the data used in our analysis in the format of R data file (.RData), and we host those data on Github[https://github.com/PangChung/ExtremePrecip], which are public accessible. However, the raw data that used to derive the temperature covariate will be only provided upon email request due to its large size. The email address is peng[dot]zhong[at]unsw[dot]edu[dot]au.

Non-publicly available data

Description

File format(s)

☐ CSV or other plain text.
☒ Software-specific binary format (.Rda, Python pickle, etc.): .Rda (.RData)
☐ Standardized binary format (e.g., netCDF, HDF5, etc.):
☐ Other (please specify):

Data dictionary

☒ Provided by authors in the following file(s):

- data/dep.fit.boot.results3.RData: Fitted results from the bootstrap scheme for the dependence model.
- data/marginal_fit_quantiles.RData: Transformed margins based on the marginal fit.
- data/temperature.RData: The derived temperature covariate over the period 1965–2020.
- data/temperature_pred.RData: The derived temperature covariate from the climate models over the period 2015–2100 under different shared socioeconomic pathways (SSP 2-4.5 or SSP 5-8.5).
- data/precip.RData: The precipitation data for the eight subregions.
- data/era5_geoinfo.RData: Shape files for the eight subregions
- data/transformed_coordinates.RData: Transformed coordinates for the eight subregions, which transformed the latitude/longitude coordinate to the Eculidean coordinate.

☐ Data file(s) is(are) self-describing (e.g., netCDF files)
☒ Available at the following URL:

- Github: https://github.com/PangChung/ExtremePrecip/

Additional Information (optional)

Part 2: Code

Abstract

Description

Code format(s)

☒ Script files
  ☒ R
  ☐ Python
  ☐ Matlab
  ☐ Other:
☒ Package
  ☒ R
  ☐ Python
  ☐ MATLAB toolbox
  ☐ Other:
☐ Reproducible report
  ☐ R Markdown
  ☐ Jupyter notebook
  ☐ Other:
☒ Shell script
☐ Other (please specify):

Supporting software requirements

Version of primary software used

Libraries and dependencies used by the code

Supporting system/hardware requirements (optional)

Parallelization used

- ☐ No parallel code used
- ☐ Multi-core parallelization on a single machine/node
  - Number of cores used:
- ☐ Multi-machine/multi-node parallelization
  - Number of nodes and cores used:

License

- ☐ MIT License (default)
- ☐ BSD
- ☐ GPL v3.0
- ☐ Creative Commons
- ☐ Other: (please specify)

Additional information (optional)

Part 3: Reproducibility workflow

Scope

The provided workflow reproduces:

- ☐ Any numbers provided in text in the paper
- ☐ The computational method(s) presented in the paper (i.e., code is provided that implements the method(s))
- ☐ All tables and figures in the paper
- ☐ Selected tables and figures in the paper, as explained and justified below:

Workflow

Location

The workflow is available:

- ☐ As part of the paper's supplementary material.
- ☐ In this Git repository:
- ☐ Other (please specify):

Format(s)

- ☐ Single master code file
- ☐ Wrapper (shell) script(s)
- ☐ Self-contained R Markdown file, Jupyter notebook, or other literate programming approach
- ☐ Text file (e.g., a readme-style file) that documents workflow
- ☐ Makefile
- ☐ Other (more detail in *Instructions* below)

Instructions

Expected run-time

Approximate time needed to reproduce the analyses on a standard desktop machine:

- ☐ < 1 minute
- ☐ 1-10 minutes
- ☐ 10-60 minutes
- ☐ 1-8 hours
- ☐ > 8 hours
- ☒ Not feasible to run on a desktop machine, as described here:

Additional information (optional)

Notes (optional)