1 Introduction

This is the data sharing module for the submitted article --- “Global models overestimate streamflow induced by rising CO_2”, and a more detailed description of the data and codes is given here for review and exchange with reviewers and interested research stakeholders.

Author: Haoshan Wei; Yongqiang Zhang*; Qi Huang; Francis H.S. Chiew; Jinkai Luan; Jun Xia; Changming Liu
Programming and Documentation: Haoshan Wei
Date: September 2023

2 Description of the program codes

2.1 Fig 1.ipynb

The plotting code for Figure 1, in which the relevant underlying data is used, is provided here, as is the plotting code for Figures 1a and b, and Figure 1c, which is pieced together on the basis of these figures; the piecing code is so simple that it is not provided here.

Figure 1a and 1b:

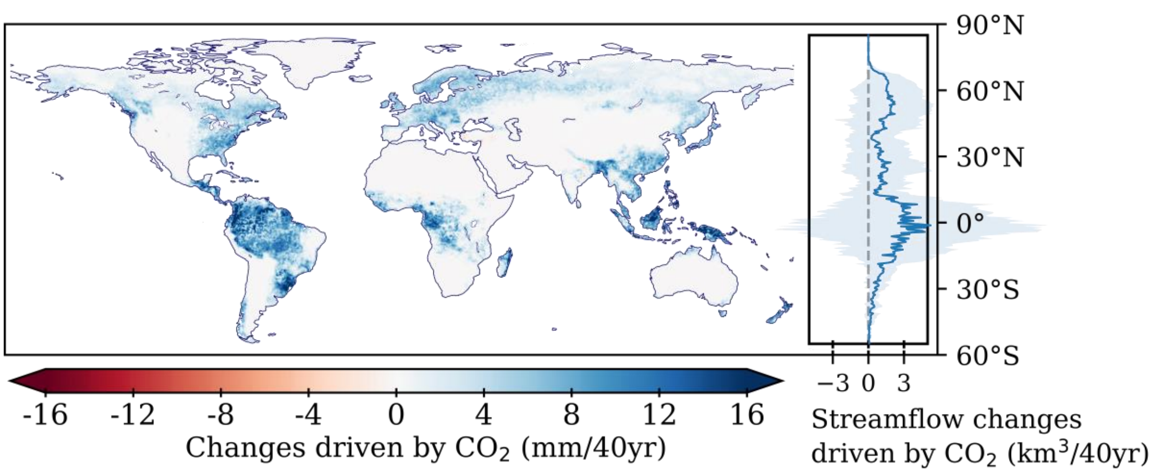
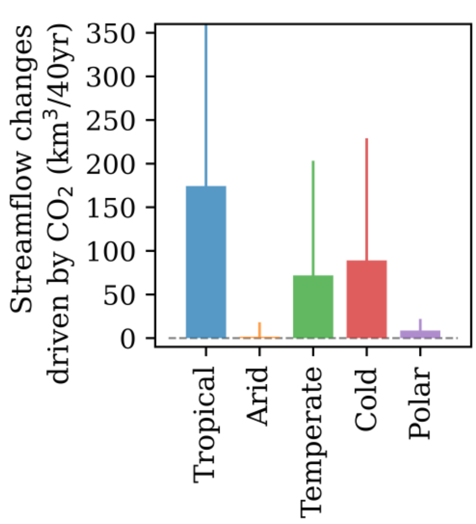


Figure 1c:



2.2 Fig 2.ipynb

The plotting code for Figure 2, in which the relevant underlying data is used, is provided here, as is the plotting code for Figure 2a and Figure 2b, which is pieced together on the basis of these figures; the piecing code is so simple that it is not provided here.

Figure 2a:

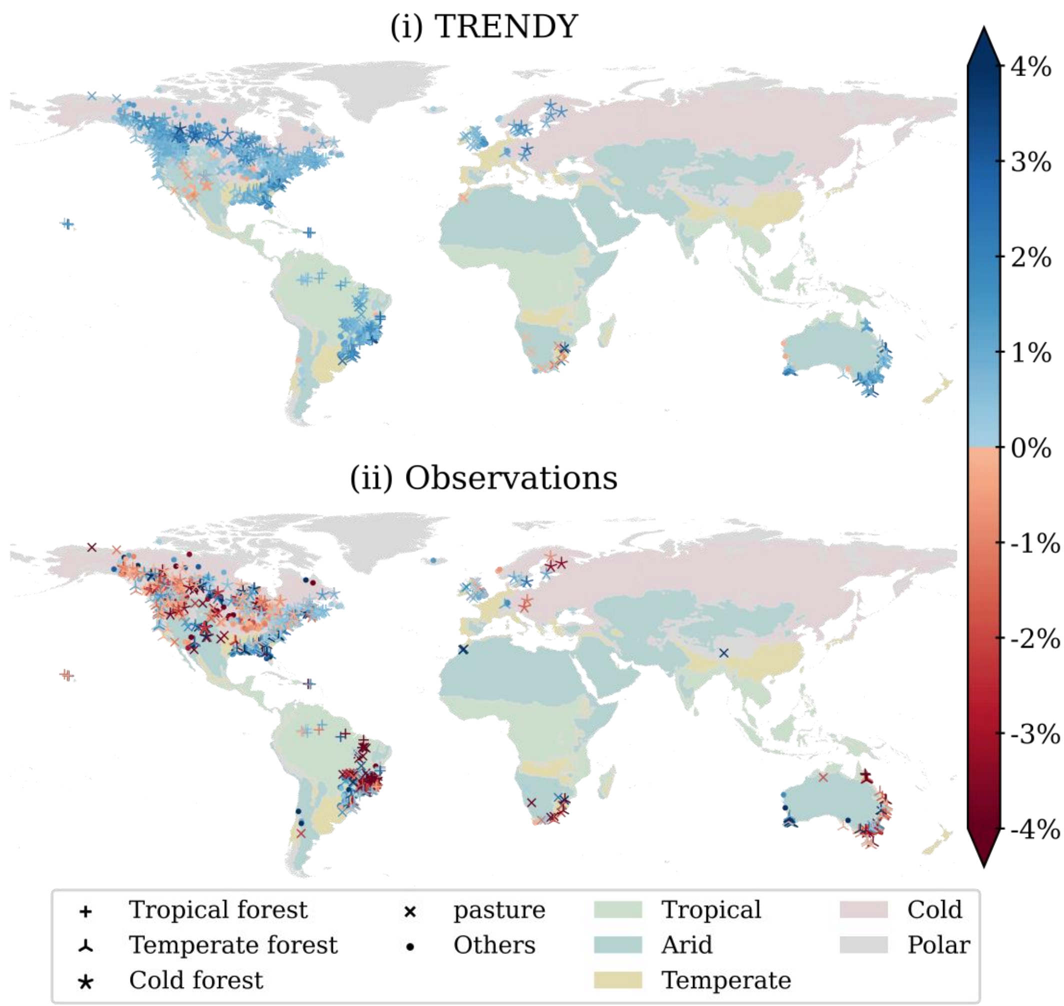
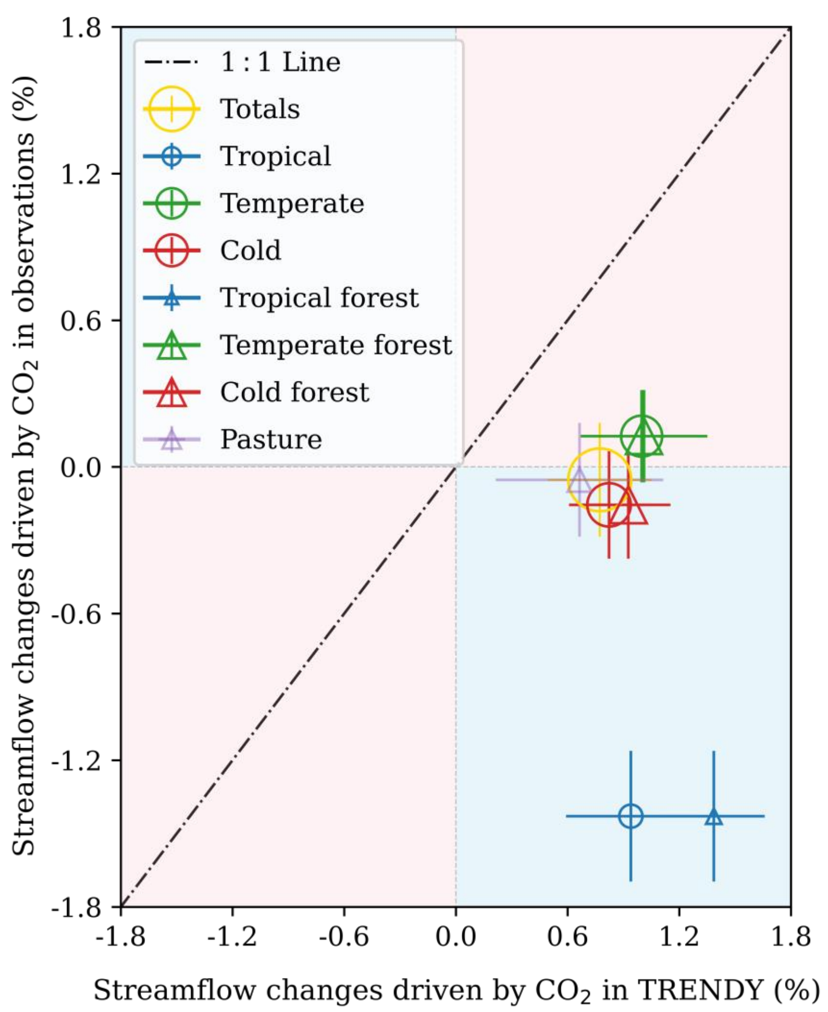
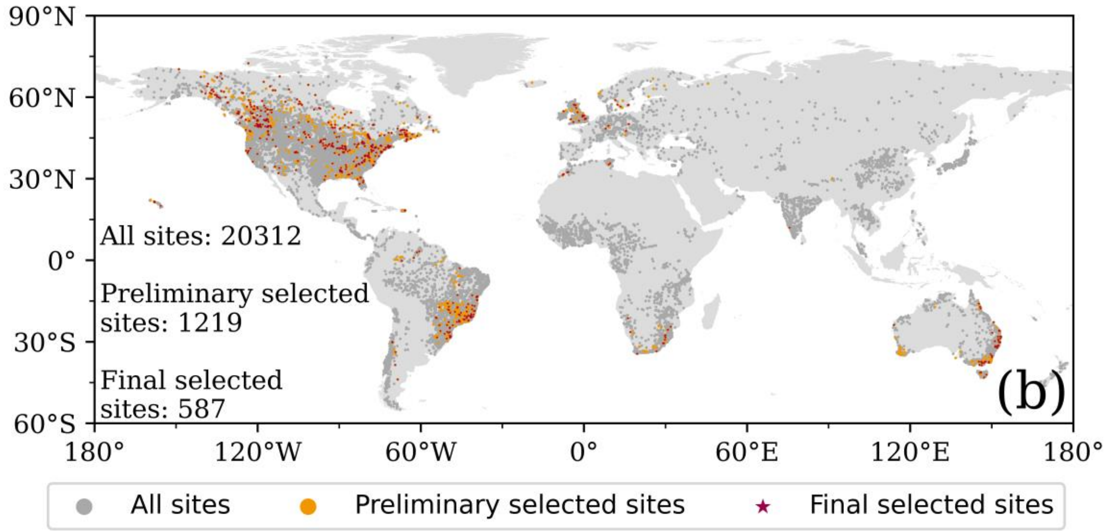

(i) TRENDY
(ii) Observations
4%
3%
2%
1%
0%
-1%
-2%
-3%
-4%
Tropical forest
Temperate forest
Cold forest
pasture
Others
Tropical
Arid
Temperate
Cold
Polar

Figure 2b:


1 : 1 Line
Totals
Tropical
Temperate
Cold
Tropical forest
Temperate forest
Cold forest
Pasture
Streamflow changes driven by CO2 in observations (%)
Streamflow changes driven by CO2 in TRENDY (%)
1.8
1.2
0.6
0.0
-0.6
-1.2
-1.8

2.3 Fig E1.ipynb

The plotting code for Figure E1 (Extended Data Fig. 1).



2.4 Fig E2.ipynb

The plotting code for Figure E2 (Extended Data Fig. 2), in which the relevant underlying data is used, is provided here, as is the plotting code for Figure E2a and Figure E2b, which is pieced together on the basis of these figures; the piecing code is so simple that it is not provided here.

Figure E2a:

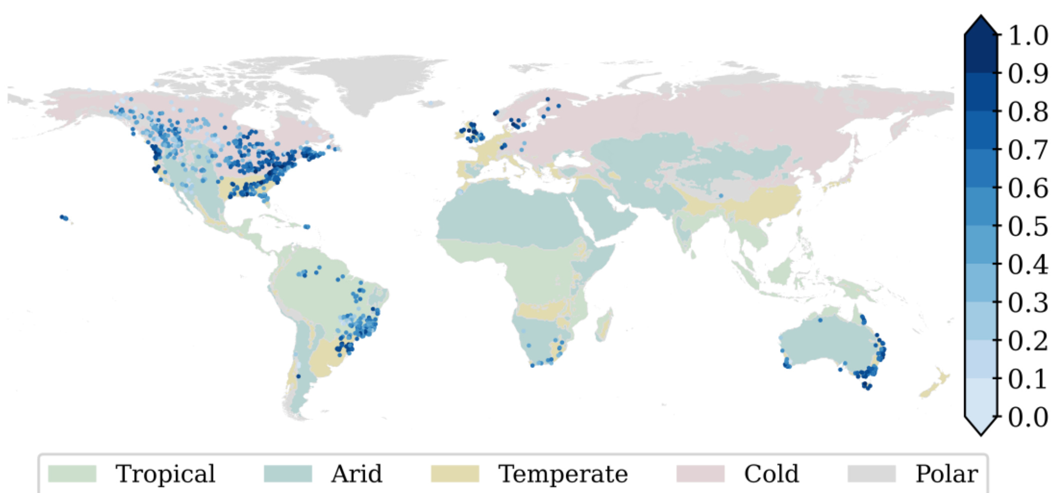
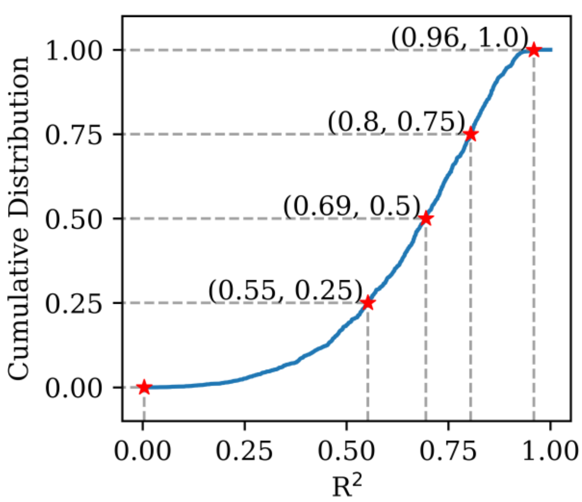


Figure E2b:



2.5 Fig E3.ipynb

The plotting code for Figure E3 (Extended Data Fig. 3), in which the relevant underlying data is used, is provided here, as is the plotting code for Figure E3a and Figure E3b, which is pieced together on the basis of these figures; the piecing code is so simple that it is not provided here.

Figure E3a:

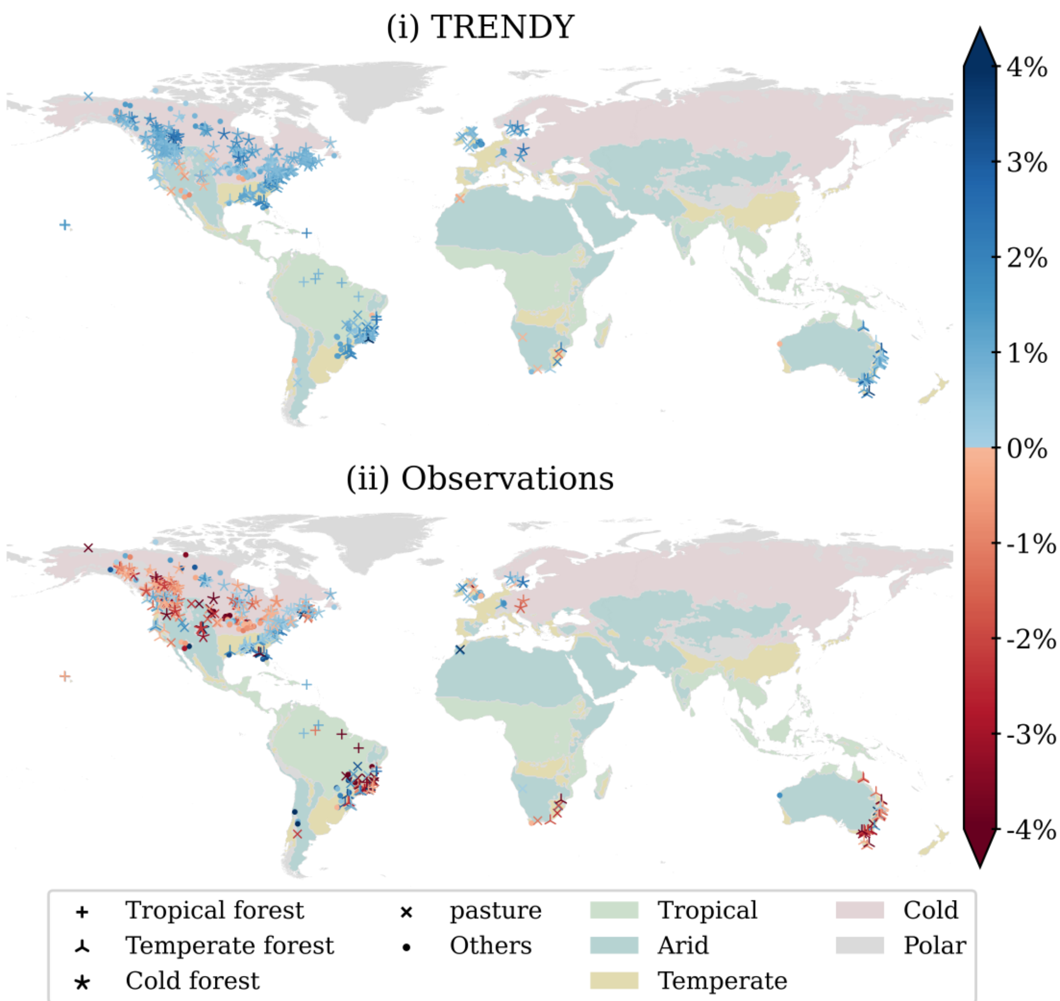
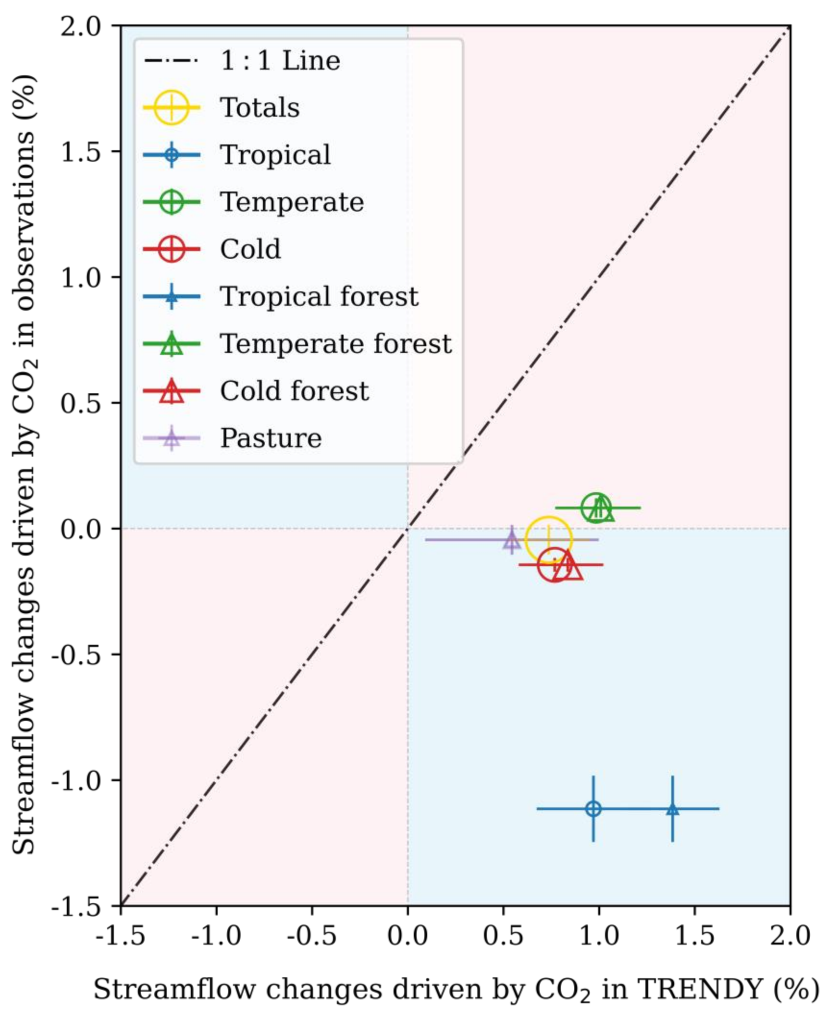


Figure E3b:



2.6 Fig S1.ipynb, Fig S2.ipynb and Fig S3 to S5.ipynb

These codes are supplemental figures to the article, and the relevant data and results used have already been provided, and the results of the run are also visible in the code. Therefore, we will not repeat it here.

3 Description of data in the folder: dataset

3.1 result.xlsx

Name	---	Name of catchments
Area	---	Area of the catchments
Lon	---	Longitude
Lat	---	Latitude
no_missing_year_len	---	Length of year with no missing daily data (1981-2020)
Index	---	Pettitt test results, 0 represents change, 1 represents no abrupt change
Others	---	"Others" in Fig. 2a
Main_climate	---	Köppen-Geiger climate classification, 0-Tropical, 1-Arid, 2-Temperate, 3-Cold, 4-Polar
ocean, urban, cropland, pasture, forest, shrubland, sparse, water	---	Average of percentages for each land use (and vegetation type)
Observations_CO2	---	Relative changes in observed streamflow driven by CO2
Trendy_CO2	---	Median of relative changes in streamflow driven by CO2 from TRENDY

Note: There are some catchments for which no official data are available, so the watershed areas of these watersheds were obtained by the catchments extraction method, which is available in the literature[1].

3.2 trendy_all_base_trend.csv

14 TRENDY models (CABLE-POP, CLASSIC, CLM5.0, DLEM, IBIS, ISAM, ISBA-CTRIP, JSBACH, JULES, LPJ-GUESS, LPX-BERN, ORCHIDEE, SDGVM, VISIT-NIES) relative changes in CO_2-driven streamflow over selected catchments. MEDIAN is the median value of these models relative changes.

3.3 observation_based_approach_result.csv

This file is the results of standardized multiple regressions for each of the 1116 catchments under the four potential evapotranspiration calculation methods, as described below:

Name	---	Name of catchments
Area	---	Area of the catchments
Lon	---	Longitude
Lat	---	Latitude

In the latter columns, the first half of the column name is an abbreviation of the method used to calculate potential evapotranspiration:

PET_FAO_	---	FAO Penman-Monteith method
PET_FAO_YANG_	---	FAO Penman-Monteith (YANG) method[2]
PET_PT_	---	Priestley-Taylor method
PET_HargreavesSamani_	---	Hargreaves-Samani method

The reference paper for FAO Penman-Monteith (YANG) is given here, and the other three methods are classical methods for calculating potential evapotranspiration, which can be found in any textbook on the subject and will not be repeated here.

The second half of the column name is an abbreviation of the specific variable:

R2	---	Goodness-of-fit R^2 for standardized regression

p_value	---	Significance coefficients for standardized regression
par_P	---	Standardized regression coefficients for precipitation (P)
par_Ep	---	Standardized regression coefficients for potential evapotranspiration (Ep)
par_CO2	---	Standardized regression coefficients for CO_2 (CO_2)
CO2_contribution	---	CO_2-driven changes in streamflow
P_contribution	---	Precipitation-driven changes in streamflow
Ep_contribution	---	Potential evapotranspiration-driven changes in streamflow
Q_change	---	Changes in streamflow
obs	---	Multi-year averages of observed streamflow

3.4 compare_CO2_have_or_not.xlsx

Standardized multiple regression with or without the inclusion of CO2 as a variable in the R^2 difference, based on the FAO Penman-Monteith (YANG) method

PET_FAO_YANG_R2_have_CO2	---	Goodness of fit with CO2 as a variable
PET_FAO_YANG_R2_no_CO2	---	Goodness of fit without CO2 as a variable

3.5 CO2_trend.nc

15 TRENDY models interpolated CO_2-driven streamflow (runoff) change values and runoff simulated change values S3 in mm/yr. 0-14 corresponds to: CABLE-POP, CLASSIC, CLM5.0, DLEM, IBIS, ISAM, ISBA-CTRIP, JSBACH, JULES, LPJ-GUESS, LPX-BERN, ORCHIDEE, SDGVM, VISIT-NIES and VISIT, respectively. Since VISIT-NIES and VISIT are extremely similar, only VISIT-NIES was retained in the calculations.

3.6 Beck_KG_5_classifications.nc

Köppen-Geiger climate classification[3] of 0.5×0.5, where, 0-Tropical, 1-Arid, 2-Temperate, 3-Cold, 4-Polar.

3.7 stations_shp

.shp data for all catchments (more than 20,000 catchments).

3.8 continent_shp

.shp data for all continents.

4 Description of data in the folder: streamflow dataset

4.1 info.xlsx

Name	---	Name of catchments
File_name	---	Filename
Area	---	Area of the catchments
Lon	---	Longitude
Lat	---	Latitude
no_missing_year_len	---	Length of year with no missing daily data (1981-2020)
LC_change[4,5]	---	Land use change (%)
irrigate_change[6]	---	Irrigated change (%)
reservoir_GDAT[7]	---	Reservoir impact (%)
reservoir_GAN[8]	---	Reservoir impact (%)
reservoir_dor_pc_pva[9]	---	Reservoir impact (%)

Note: Reservoir impact is calculated as reservoir water capacity/average multi-year streamflow. Here, the selected catchments have no reservoirs, or the reservoir impact is 0.

4.2 Yearly_streamflow_dataset

The folder contains detailed annual change data for 1116 sites, with the file name being the catchment name. Here, specific data is included:

Year --- Years from 1981 to 2020
Flow_mm --- Annual streamflow for the year
P_MSWEP --- Calculated annual precipitation[10]
CO2 --- Annual value of carbon dioxide

Annual atmospheric CO_2 concentration data calculated from monthly data; monthly atmospheric CO_2 concentration data were obtained from NOAA globally averaged marine surface monthly mean data (ftp://aftp.cmdl.noaa.gov/products/trends/co2/co2 mm gl.txt)

Annual potential evapotranspiration was calculated by these 4 methods:

PET_FAO --- FAO Penman-Monteith method
PET_FAO_YANG_ --- FAO Penman-Monteith (YANG) method[2]
PET_PT_ --- Priestley-Taylor method
PET_HargreavesSamani --- Hargreaves-Samani method

The data needed to calculate potential evapotranspiration were obtained from the literature[11].

The percentage of land use change per year in the catchment is also given, including: ocean, urban, cropland, pasture, forest, shrubland, sparse and water[4,5].

References

1. Xie, J., Liu, X., Bai, P. & Liu, C. Rapid Watershed Delineation Using an Automatic Outlet Relocation Algorithm. *Water Resources Research* **58**, (2022).
2. Yang, Y., Roderick, M. L., Zhang, S., McVicar, T. R. & Donohue, R. J. Hydrologic implications of vegetation response to elevated CO2 in climate projections. *Nature Clim Change* **9**, 44–48 (2019).
3. Beck, H. E. *et al.* Present and future Köppen-Geiger climate classification maps at 1-km resolution. *Sci Data* **5**, 180214 (2018).
4. Winkler, K., Fuchs, R., Rounsevell, M. & Herold, M. Global land use changes are four times greater than previously estimated. *Nat Commun* **12**, 2501 (2021).
5. Winkler, K., Fuchs, R., Rounsevell, M. D. A. & Herold, M. HILDA+ Global Land Use Change between 1960 and 2019. (2020) doi:10.1594/PANGAEA.921846.
6. Global Map of Irrigation Areas (GMIA) | Land & Water | Food and Agriculture Organization of the United Nations | Land & Water | Food and Agriculture Organization of the United Nations. https://www.fao.org/land-water/land/land-governance/land-resources-planning-toolbox/category/details/en/c/1029519/.
7. Zhang, A. T. & Gu, V. X. Global Dam Tracker: A database of more than 35,000 dams with location, catchment, and attribute information. *Sci Data* **10**, 111 (2023).
8. Lehner, B. Global Reservoir and Dam (GRanD) database.
9. Linke, S. *et al.* Global hydro-environmental sub-basin and river reach characteristics at high spatial resolution. *Sci Data* **6**, 283 (2019).
10. Beck, H. E. *et al.* MSWEP V2 Global 3-Hourly 0.1° Precipitation: Methodology and Quantitative Assessment. *Bulletin of the American Meteorological Society* **100**, 473–500 (2019).
11. Beck, H. E. *et al.* MSWX: Global 3-Hourly 0.1° Bias-Corrected Meteorological Data Including Near-Real-Time Updates and Forecast Ensembles. *Bulletin of the American Meteorological Society* **103**, E710–E732 (2022).