Supplementary information for

Machine learning corrects climate model biases in global flood projections

Boen Zhang[1,*], Louise Slater[1], Simon Moulds[2], Michel Wortmann[3], Neil Hart[1], Yinxue Liu[1,4], Jiabo Yin[5], Xihui Gu[6], Grey Nearing[7], Manuela Brunner[8,9,10], Stephen Roberts[11]

[1] School of Geography and the Environment, University of Oxford, Oxford, UK.

[2] School of Geosciences, University of Edinburgh, Edinburgh, UK

[3] European Centre for Medium-Range Weather Forecasts, Shinfield Park, Reading, UK.

[4] Geography and Environment, Loughborough University, Loughborough, UK

[5] State Key Laboratory of Water Resources Engineering and Management, Wuhan University, Wuhan, PR China

[6] Department of Atmospheric Science, School of Environmental Studies, China University of Geosciences, Wuhan 430074, China

[7] Google Research, Mountain View, CA, USA

[8] Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland

[9] Institute for Snow and Avalanche Research SLF, Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Davos Dorf, Switzerland

[10] Climate Change, Extremes and Natural Hazards in Alpine Regions Research Center CERC, Davos Dorf, Switzerland

[11] Department of Engineering Science, University of Oxford, Oxford, United Kingdom

*boen.zhang@ouce.ox.ac.uk

This document contains Supplementary Text S1, Supplementary Tables S1–5, and Supplementary Figures S1–16.

Text S1. Model performance metrics.

We use the Nash-Sutcliffe Efficiency (NSE), the Normalized Root Mean Squared Error (nRMSE), and the Kling–Gupta efficiency (KGE) metrics, along with the three components of KGE (i.e., Pearson's correlation, bias ratio, and variability ratio) to evaluate the performance of random forest models.

The NSE is expressed as

$$NSE = 1 - \frac{\sum(O_i - S_i)^2}{\sum(O_i - \bar{O})^2} \tag{1}$$

where S_i and O_i denote simulations and observations, respectively, and are applicable to all the following equations; $\bar{O}$ is the observed mean.

The nRMSE is expressed as

$$nRMSE = \frac{1}{O_{max} - O_{min}} \sqrt{\frac{\sum(S_i - O_i)^2}{N}} \tag{2}$$

where N denotes the number of samples. O_{max} and $O_{\min}$ denote the maximum and minimum observations, respectively.

The KGE is expressed as

$$KGE = 1 - \sqrt{(r-1)^2 + (\beta - 1)^2 + (\gamma - 1)^2} \tag{3}$$

where r, β, and γ represent Pearson's correlation, bias ratio, and variability ratio, respectively, and are expressed as

$$r = \frac{\sum(S_i - \bar{S})(O_i - \bar{O})}{\sqrt{\sum(S_i - \bar{S})^2 \sum(O_i - \bar{O})^2}} \tag{4}$$

$$\beta = \bar{S}/\bar{O} \tag{5}$$

$$\gamma = \frac{\sigma_S/\bar{S}}{\sigma_O/\bar{O}} \tag{6}$$

In perfect prediction, where KGE and NSE equal unity, the $r = \beta = \gamma = 1$ and nRMSE = 0. According to equations 5 and 6, $\beta < 1$ corresponds with an overall negative bias, in which $\bar{S} < \bar{O}$. Similarly, $\beta > 1$ indicates a positive bias in prediction.

Table S1. Existing river flood projection studies. CMIP5/6 = Coupled Model Intercomparison Project Phase 5/6. ISIMIP = Inter-Sectoral Impact Model Intercomparison Project. CORDEX = Coordinated Regional Climate Downscaling Experiment. HM = hydrologic model. VIC = Variable Infiltration Capacity hydrologic model. CPM = convection-permitting climate model. AMS = annual maximum streamflow. MAX/MAX7 = maximum daily/7-day streamflow. Q90 = 90th percentile of daily streamflow.

No.	Approach	Index	Area
1	CMIP6 + ISIMIP + LSTM	AMS	China[1]
2	CMIP5/CMIP6/CORDEX + HM	100-year flood	Europe[2]
3	CMIP6 + CaMa-Flood	100-year flood	Global[3]
4	ISIMIP + VIC + CaMa-Flood	100-year flood	China urban areas[4]
5	ISIMIP + LSTM	MAX/MAX7/Q90	Global 2062 river basins[5]
6	ISIMIP	MAX7	Global[6]
7	CMIP5 runoff + CaMa-Flood		Global[7]
8	ISIMIP	River flow seasonality	Global[8]
9	EC-EARTH3-HR + CMIP5 + LISFLOOD + local value of the flood protection	Six return periods (10-500 years)	Global[9]
10	CMIP6 + VIC	Widespread flooding (streamflow > Q99 for all subbasins within a river basin)	India[10]
11	Bias-corrected CMIP5 + VIC	1 % annual exceedance probability	USA[11]
12	Bias-corrected CMIP5 + NOAA's HL-RDHM + LISFLOOD-FP	100-year flood peak	Pennsylvania, USA[12]
13	Bias-corrected CMIP5 + HBV + LISFLOOD-FP	100-year flood	USA[13]
14	HAPPI GCM + Tropical cyclone rainfall model + LISFLOOD-FP	Hurricane rainfall-induced flood	Puerto Rico[14]
15	CPM + LISFLOOD-FP	Urban flooding	UK[15]
16	CMIP6 + GAMLSS	Annual exceedance probabilities (e.g., 1 in 10)	USA[16]
17	Bias-corrected CMIP5 + HM	Annual exceedance probabilities (e.g., 1 in 10)	Australia[17]
18	CMIP6 runoff – base flow separation	AMS	Global[18]
19	Downscaled CMIP6 + GAMLSS	10-year, 50-year, and 100-year flood	West-Central Himalayas[19]

Table S2. 19 CMIP6 models used in this study

No.	Model	Variant label	Horizontal resolution (lon × lat)
1	ACCESS-CM2[20]	r1i1p1f1	1.9° × 1.3°
2	ACCESS-ESM1-5[21]	r1i1p1f1	1.9° × 1.2°
3	CanESM5[22]	r1i1p1f1	2.8° × 2.8°
4	CNRM-ESM2-1[23]	r1i1p1f2	1.4° × 1.4°
5	CNRM-CM6-1[24]	r1i1p1f2	1.4° × 1.4°
6	AWI-ESM-1-REcoM[25]	r1i1p1f1	1.9° × 1.9°
7	EC-Earth3[26]	r1i1p1f1	0.7° × 0.7°
8	EC-Earth3-Veg[27]	r1i1p1f1	0.7° × 0.7°
9	GFDL-ESM4[28]	r1i1p1f1	1.3° × 1°
10	INM-CM5-0[29]	r1i1p1f1	2° × 1.5°
11	HadGEM3-GC31-LL[30]	r1i1p1f3	1.9° × 1.3°
12	MIROC6[31]	r1i1p1f1	1.4° × 1.4°
13	MIROC-ES2L[32]	r1i1p1f2	2.8° × 2.8°
14	MPI-ESM1-2-HR[33]	r1i1p1f1	0.9° × 0.9°
15	MPI-ESM1-2-LR[34]	r1i1p1f1	1.9° × 1.9°
16	MRI-ESM2-0[35]	r1i1p1f1	1.1° × 1.1°
17	NorESM2-LM[36]	r1i1p1f1	2.5° × 1.9°
18	UKESM1-0-LL[37]	r1i1p1f2	1.9° × 1.3°
19	IPSL-CM6A-LR[38]	r1i1p1f1	2.5° × 1.3°

Table S3. List of climate predictors used in the ML models. Area-weighted average values are extracted for each catchment, variable, and year, and then annual mean (mean) and standard deviation (std) are calculated for each variable across all years in the dataset.

No.	Variable	Unit	Potential predictors
1	Total precipitation (Psum)	mm	Psum_mean, Psum_std
2	Maximum 1-day precipitation (Rx1day)	mm	Rx1day_mean, Rx1day_std
3	Maximum 1-month precipitation (Rx1mon)	mm	Rx1mon_mean, Rx1mon_std
4	Maximum total precipitation of consecutive wet days (Pvol)	mm	Pvol_mean, Pvol_std
5	Maximum 5-day precipitation (Rx5day)	mm	Rx5day_mean, Rx5day_std
6	Relative humidity (RH)	%	RH_mean
6	Average snow melting (SMave)	mm	SMave_mean
7	Maximum near-surface air temperature (Tmax)	degreeC	Tmax_mean
8	Minimum near-surface air temperature (Tmin)	degreeC	Tmin_mean
9	Maximum snow melting (SMmax)	mm	SMmax_mean
10	Surface downwelling shortwave radiation (RSDS)	W/m^2	RSDS_mean
11	Wind speed (WS)	m/s	WS_mean
12	Surface snow amount (SNW)	kg/m^2	SNW_mean

Table S4. Geographic predictors used in the ML models

No.	Predictor class	Predictor	Unit	Data source	Extraction method
1	Land cover	Fractional built-up area	%	ESACCI land cover map in the year 2000[39]	Land cover fraction
2	Land cover	Fractional forest	%		
3	Land cover	Fractional grassland	%		
4	Geomorphology	Slope	–	HydroSHEDS v1 DEM[40]	Area-weighted average
5	Geomorphology	Elevation	m		
6	Soil	Thickness of the permeable layer	m	Global 1-km Gridded Thickness of Soil, Regolith, and Sedimentary Deposit Layers[41]	
7	Soil	Permeability	Log(k)	GLHYMPS 2.0[42]	
8	Climate	Aridity index	–	Global Aridity Index and Potential Evapotranspiration Database - Version 3[43]	
9	Climate	Potential evapotransipiration	mm		
10	Climate	Fractional arid climate zone	%	Köppen-Geiger climate classification at an 1-km resolution for the present-day (1980–2016)[44]	Climate zone fraction
11	Climate	Fractional cold climate zone	%		
12	Climate	Fractional temperate climate zone	%		
13	Climate	Fractional tropical climate zone	%		
14	Hydrologic signature	Runoff ratio	–	MSWEP v2.2 rainfall[45] and GloFAS v3.1 river discharge[46]	Area-weighted average rainfall is extracted; River discharge is extracted for the grid cell that is closest to the outlet point.
15	Hydrologic signature	Streamflow precipitation elasticity	–		
16	Hydrologic signature	Baseflow index	–		
17	Hydrologic signature	Richards-Baker flashiness index[47]	–		
18	Hydrologic signature	Groundwater storage (Q10/Q50);	–		
19	Hydrologic signature	High-flow frequency	day		
20	Hydrologic signature	Coefficient of streamflow variation	–		
21	Coordinate	Longitude	–	–	–
22	Coordinate	Latitude	–	–	–

23	Human activity	Population density	%	WorldPop[48]	Area-weighted average
24	Human activity	GDP per capita	$	Gridded global datasets for Gross Domestic Product in 2000[49]	Area-weighted average
25	Human activity	Number of dams	%	Global Dam Tracker[50]	Number of upstream dams

Table S5. Data sources of global streamflow observations

No.	Data source	URL
1	Global Runoff Data Centre (GRDC)	https://www.bafg.de/GRDC/EN/Home/homepage_node.html
2	The Bureau of Meteorology of the Australian Government (BOM)	http://www.bom.gov.au/
3	The Brazilian National Water Agency and Bruno Guimarães for his Python webscraping code "scraping-hidroweb" (available on GitHub)	http://www.snirh.gov.br/hidroweb/apresentacao
4	The Spanish Ecological transition ministry	https://ceh.cedex.es/anuarioaforos/default.asp
5	Eaufrance, the public water information service in France	https://www.eaufrance.fr/
6	National Water Data Archive of Canda (HYDAT)	https://collaboration.cmc.ec.gc.ca/cmc/hydrometrics/www/
7	Polish Institute of Meteorology and Water Management (IMGW)	https://www.imgw.pl/
8	The National River Flow Archive (NRFA), based at the UK Centre for Ecology & Hydrology	https://nrfa.ceh.ac.uk/
9	The Norwegian Water Resources and Energy Directorate (NVE)	https://www.nve.no/english/
10	The U.S. Geological Survey	https://waterdata.usgs.gov/nwis
11	Environmental Information System for water resources and modelling for Africa (SIEREM)	http://www.hydrosciences.fr/sierem/
12	Colombia Institute of Hydrology, Meteorology and Environmental Studies	http://institucional.ideam.gov.co/jsp/index.jsf
13	Federal Office for the Environment (Switzerland)	https://www.bafu.admin.ch/bafu/en/home.html
14	The Hydroinformatic Data Center managed by Ministry of Water Resources of China	http://xxfb.hydroinfo.gov.cn/ssIndex.html
15	Slovenian Environment Agency	https://www.gov.si/en/state-authorities/bodies-within-ministries/slovenian-environment-agency/
16	India- Water Resource Information System (WRIS)	https://indiawris.gov.in/wris/#/

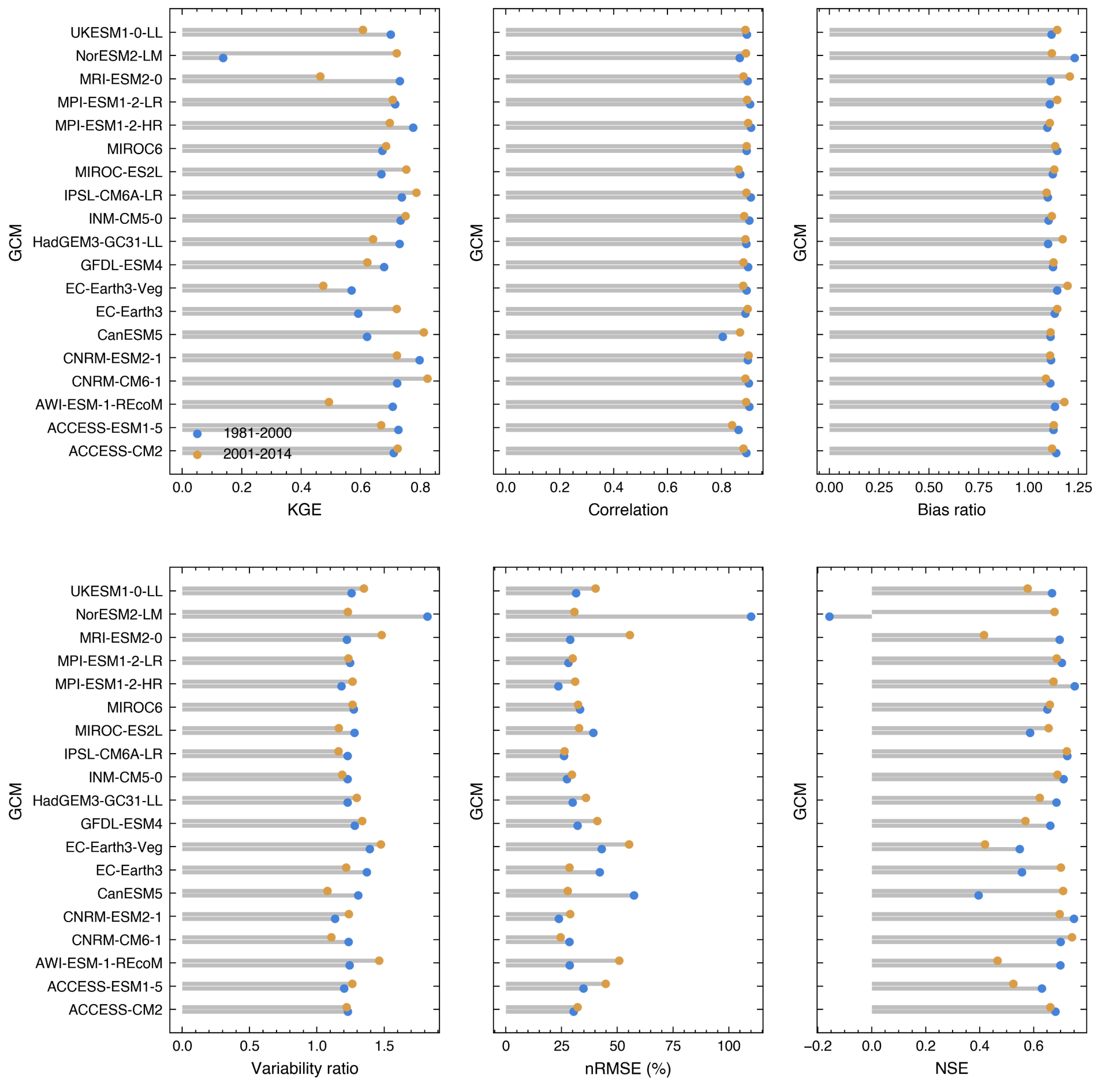


Fig. S1 Full cross-product cross-validation test of the ML models. The climate predictors and the 1-in-10-year flood magnitudes calculated for the 1981−2000 and 2001−2014 periods, respectively. A total of 20 cross-validation splits (i.e., 10 random spatial splits and 2 temporal splits) are generated to conduct the test.

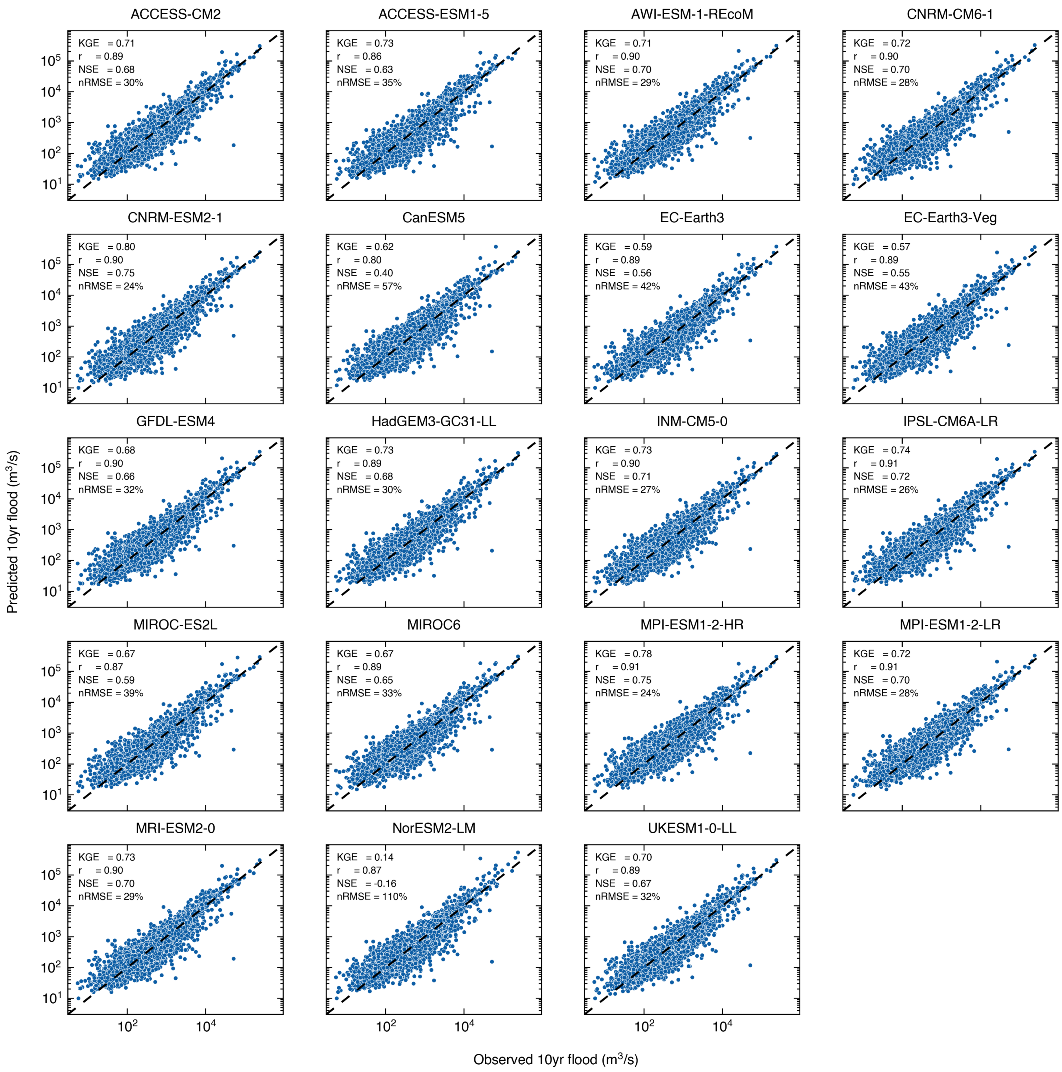


Fig. S2 Same as Fig. S1 but shows cross-validated predictions during the period 1981–2000.

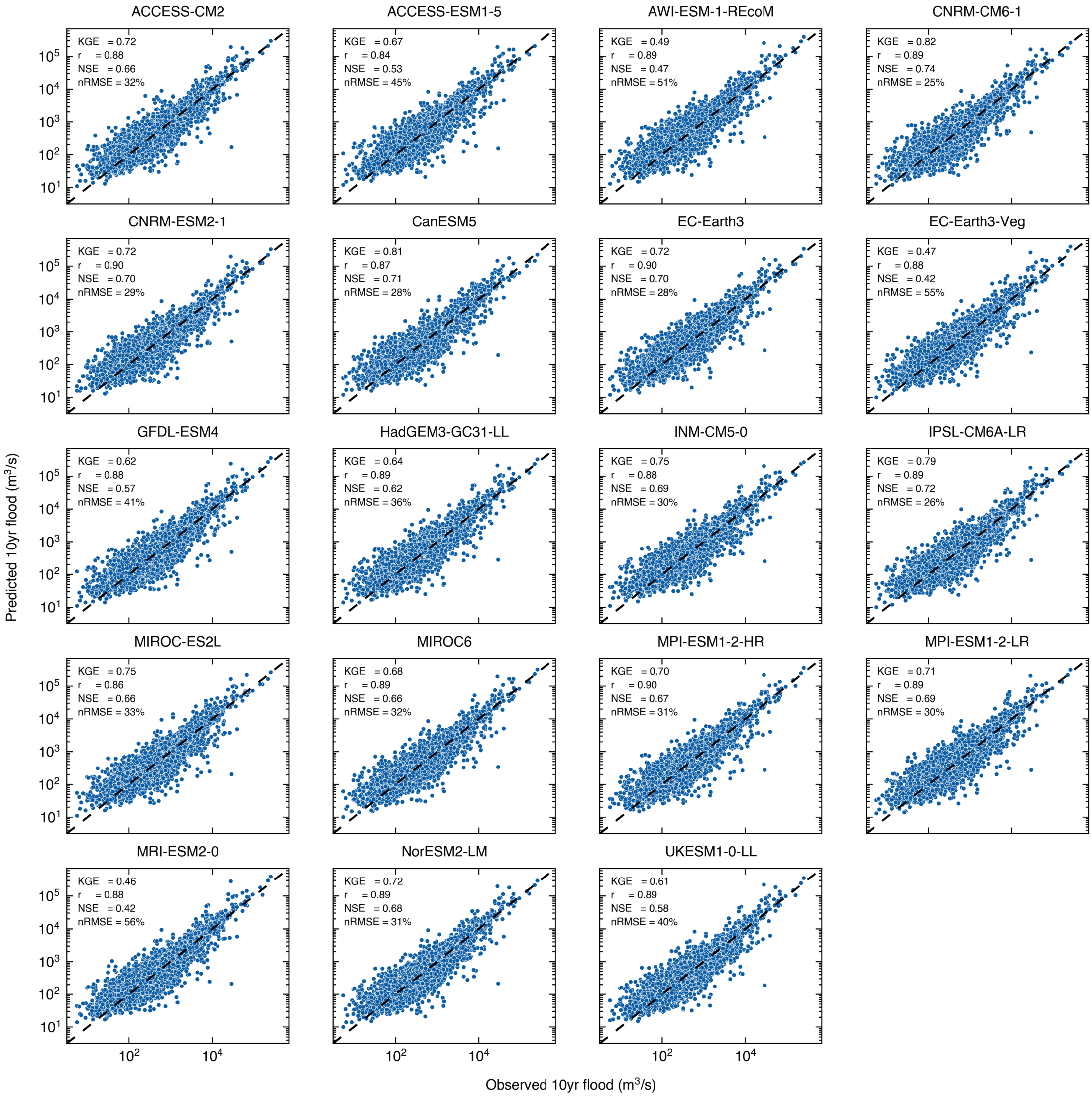


Fig. S3 Same as Fig. S1 but shows cross-validated predictions during the period 2000–2014.

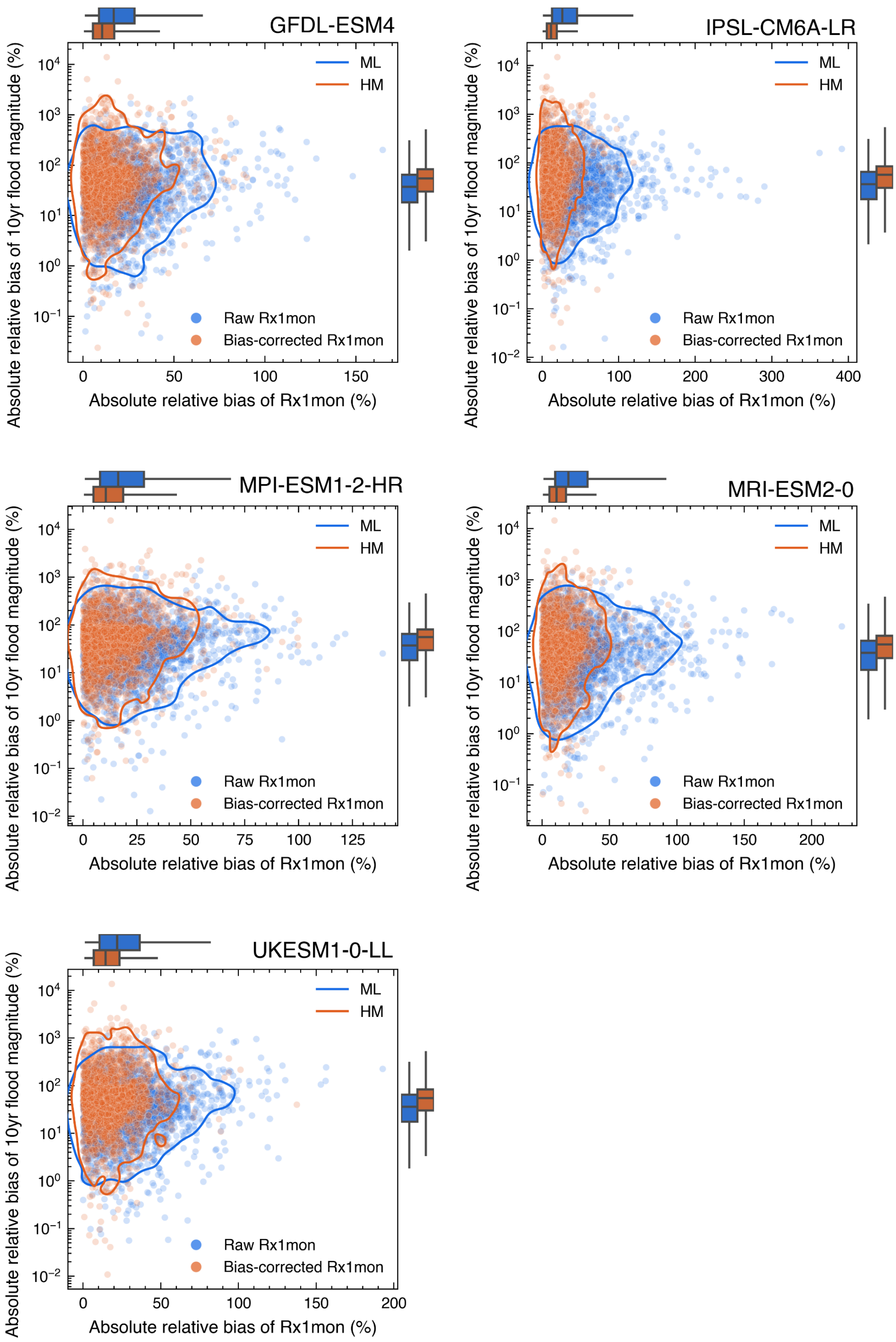


Fig. S4 Bias of CMIP6 annual maximum 1-month precipitation (Rx1mon) and of 10-year flood magnitude generated by the ensemble mean ML (blue) and HM (orange) predictions; contours represent the 95% confidence contour, indicating that 95% of the probability mass lies within the contour. Boxplots alongside each subplot indicate the distribution of flood biases and Rx1mon biases, respectively, and are significantly different ($P < 0.01$) based on the paired t-test. The boxplots whisker extend to the 2.5% and 97.5% percentiles.

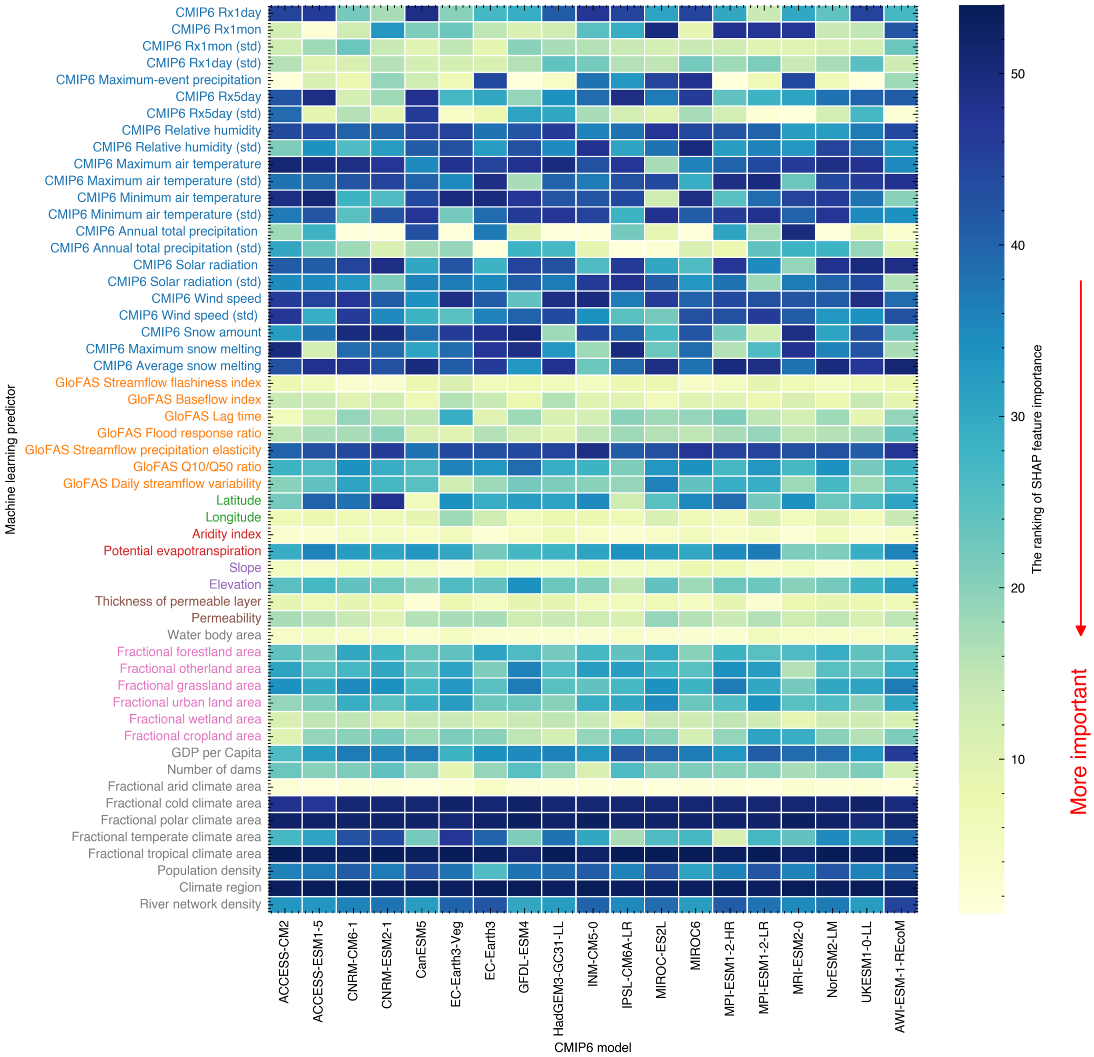


Fig. S5 Feature importance based on SHAP values. The mean absolute SHAP values for each feature were sorted to obtain the global feature importance and the procedure was repeated for each of the 19 CMIP6 models. The ranking from 1 to 54 indicates decreasing feature importance.

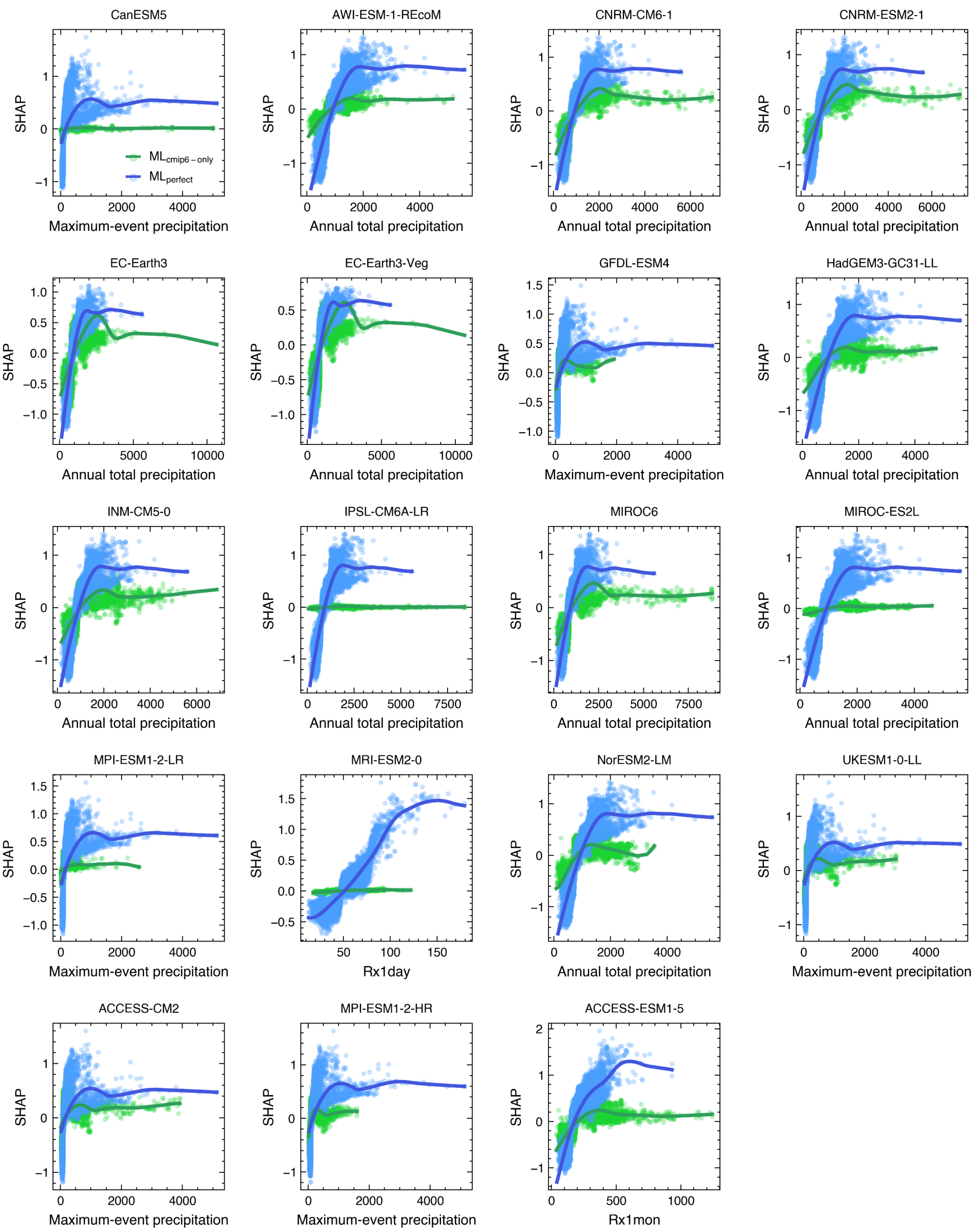


Fig. S6 We compare a simulation-only ML model ("CMIP6 only", $ML_{cmip6\text{-}only}$, shown in green) trained on climate model simulations and a perfect prediction ML model ($ML_{perfect}$, shown in blue, as a baseline) in which the most important CMIP6 precipitation predictor (e.g., annual total precipitation, or another variable) is replaced by the equivalent predictor calculated from the MSWEP dataset. The most important climate predictor in each climate model is identified through the SHapley Additive exPlanations (SHAP) algorithm (see Fig. S5). The two ML models are trained separately over the period 1981−2014. We compare the contribution of the precipitation predictor to the flood magnitude in the two ML models using SHAP values. Positive SHAP values mean that the predictor affects the flood magnitude positively, and vice versa. The higher the absolute SHAP value, the more the predictor contributes to the flood magnitude. Each coloured circle represents a gauging station. The smooth lines were fitted with LOESS (locally weighted scatterplot smoothing) regression to highlight the average differences between the two ML models.

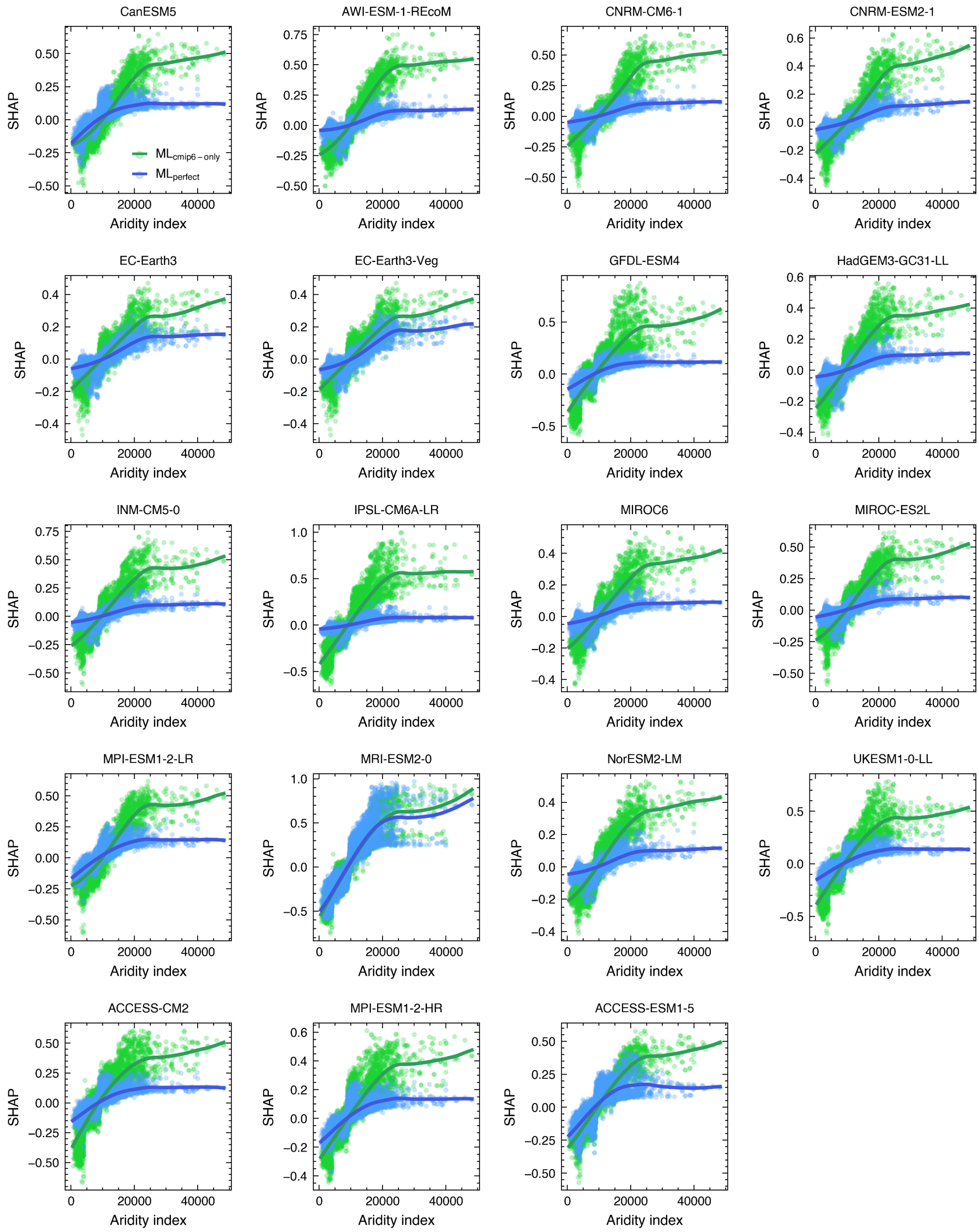


Fig. S7 Same as Fig. S6 but the contribution of the aridity index to the flood magnitude in the two ML models (*$ML_{cmip6\text{-}only}$* and *$ML_{perfect}$*) are compared.

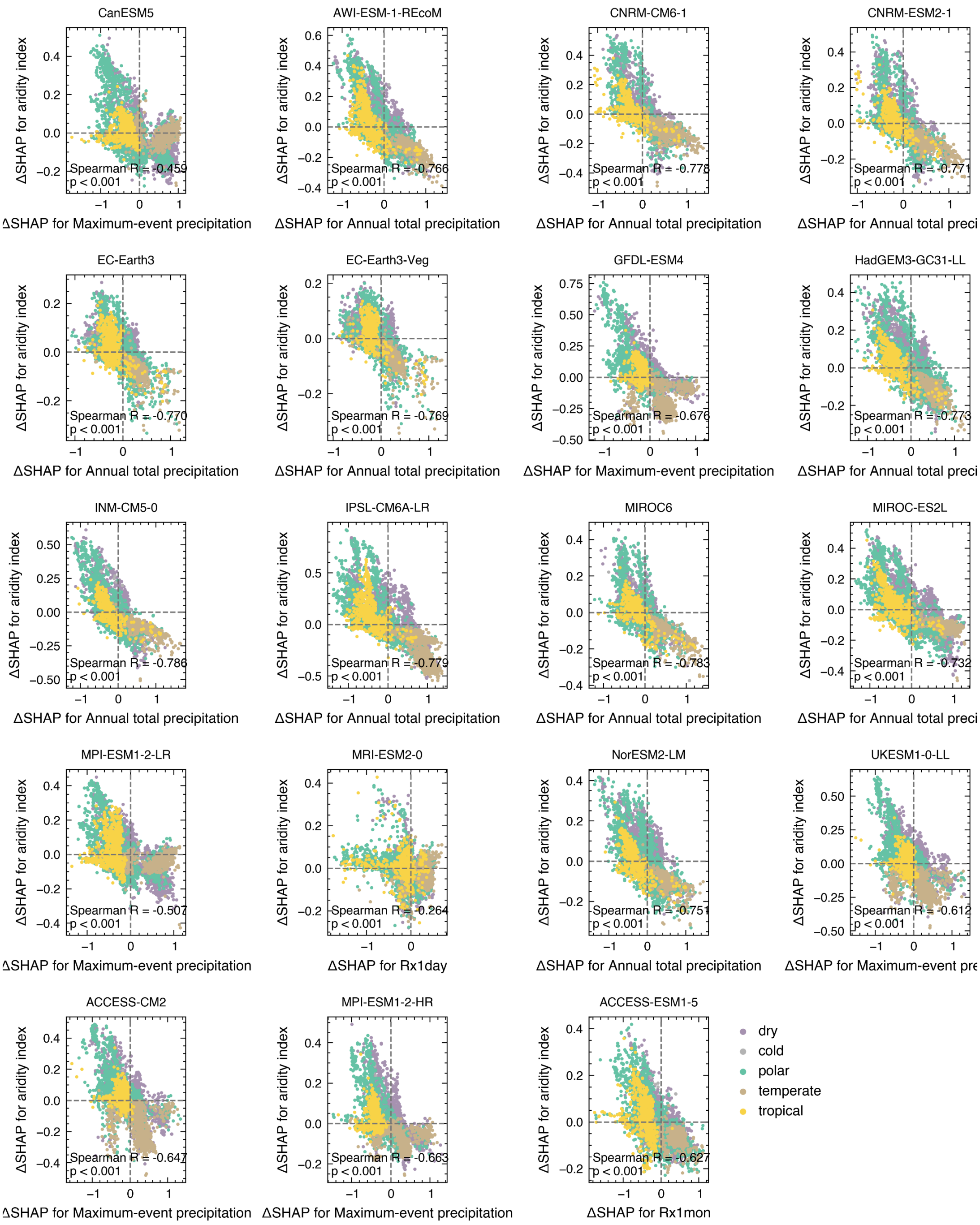


Fig. S8 The difference in SHAP values (i.e., ΔSHAP) of the most important climate variable for the $ML_{cmip6\text{-}only}$ and $ML_{perfect}$ models is compared with that of the aridity index (y-axis). It should be noted that the most important climate variable is identified through the SHapley Additive exPlanations (SHAP) algorithm for the $ML_{cmip6\text{-}only}$ model (see Fig. S5), while the most important climate variables for the $ML_{cmip6\text{-}only}$ and $ML_{perfect}$ models may be different. Positive ΔSHAP values mean that the $ML_{cmip6\text{-}only}$ model increases the contribution of the predictor to flooding compared to the $ML_{perfect}$ model. The Spearman correlation and associated p-value are labelled in the bottom left of each panel. It should be noted that the most important climate predictor (x-axis) shown here varies with the climate model.

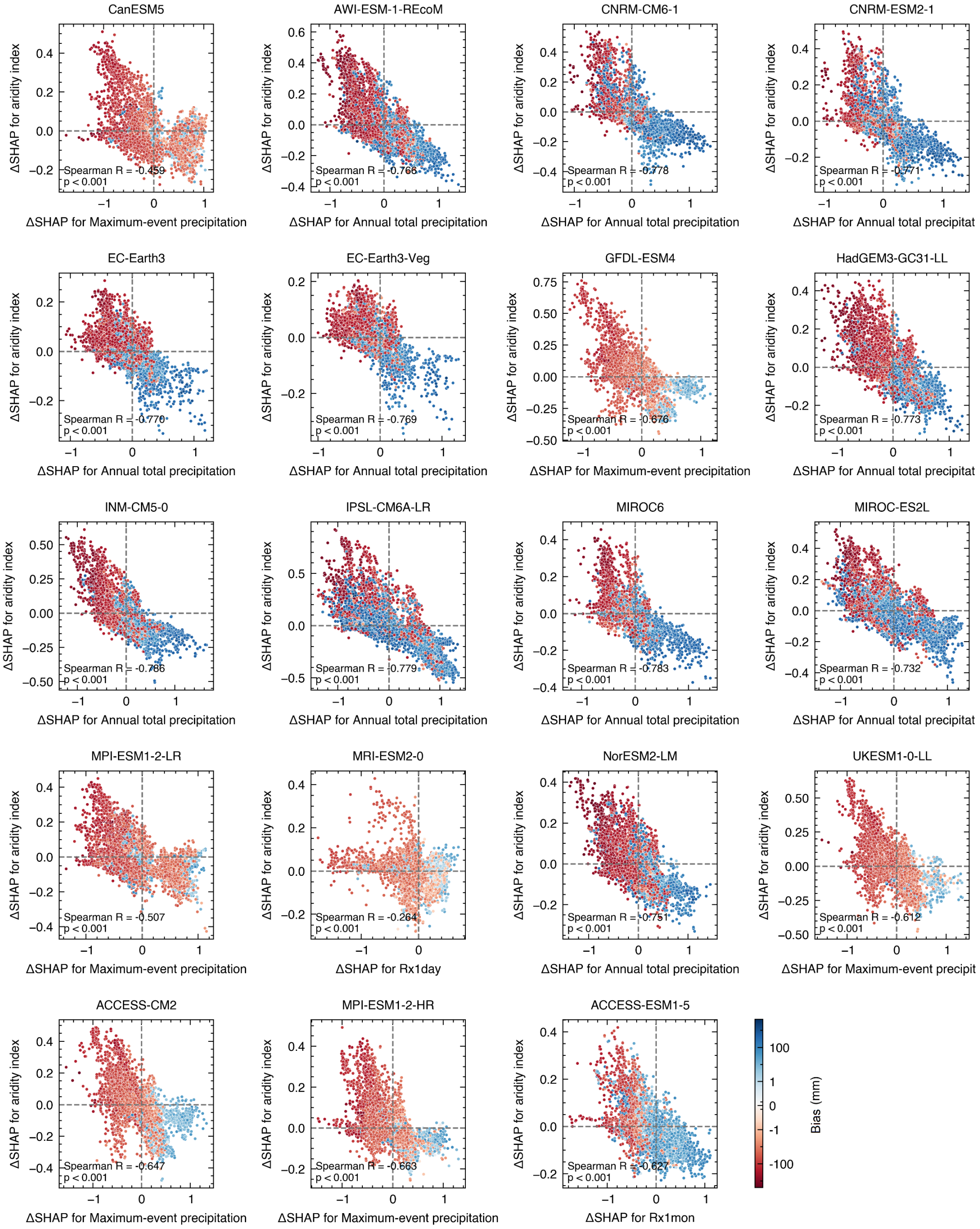


Fig. S9 Same as Fig. S8 but the dots are coloured by the bias of the climate predictors. It should be noted that the most important climate predictor (x-axis) shown here varies with the climate model (each panel is a different ML model and climate model).

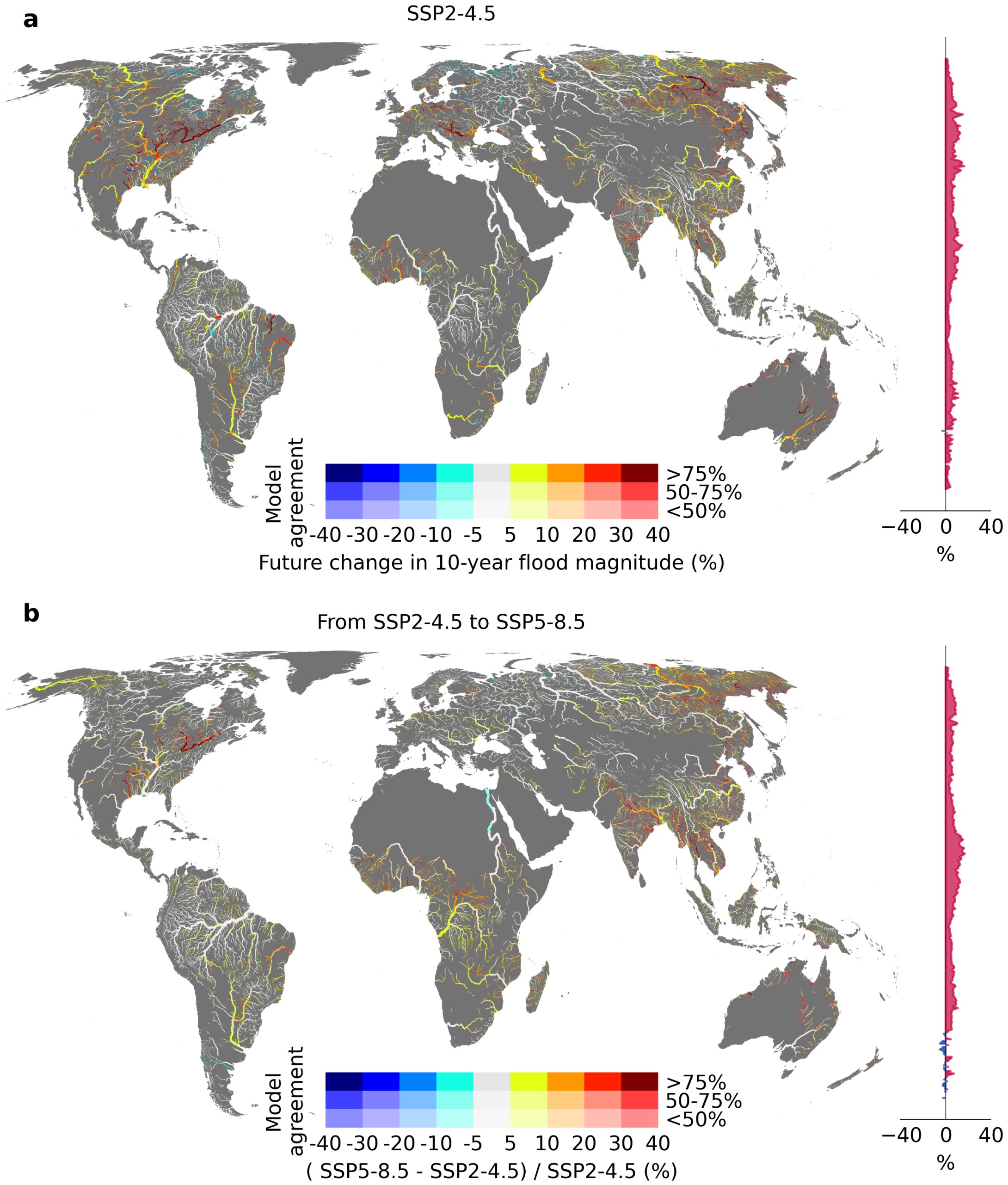


Fig. S10 Multi-model ensemble mean in the projected changes of flood magnitude between historical (1981−2014) and future (2071−2100) periods. **a** indicates future changes under SSP2-4.5 and **b** indicates the difference between SSP2-4.5 and SSP5-8.5. The colour in the maps shows the magnitude of change and the saturation indicates the agreement, among ensemble members, in the sign of change. The inset graph on the right shows the latitudinal mean flood change.

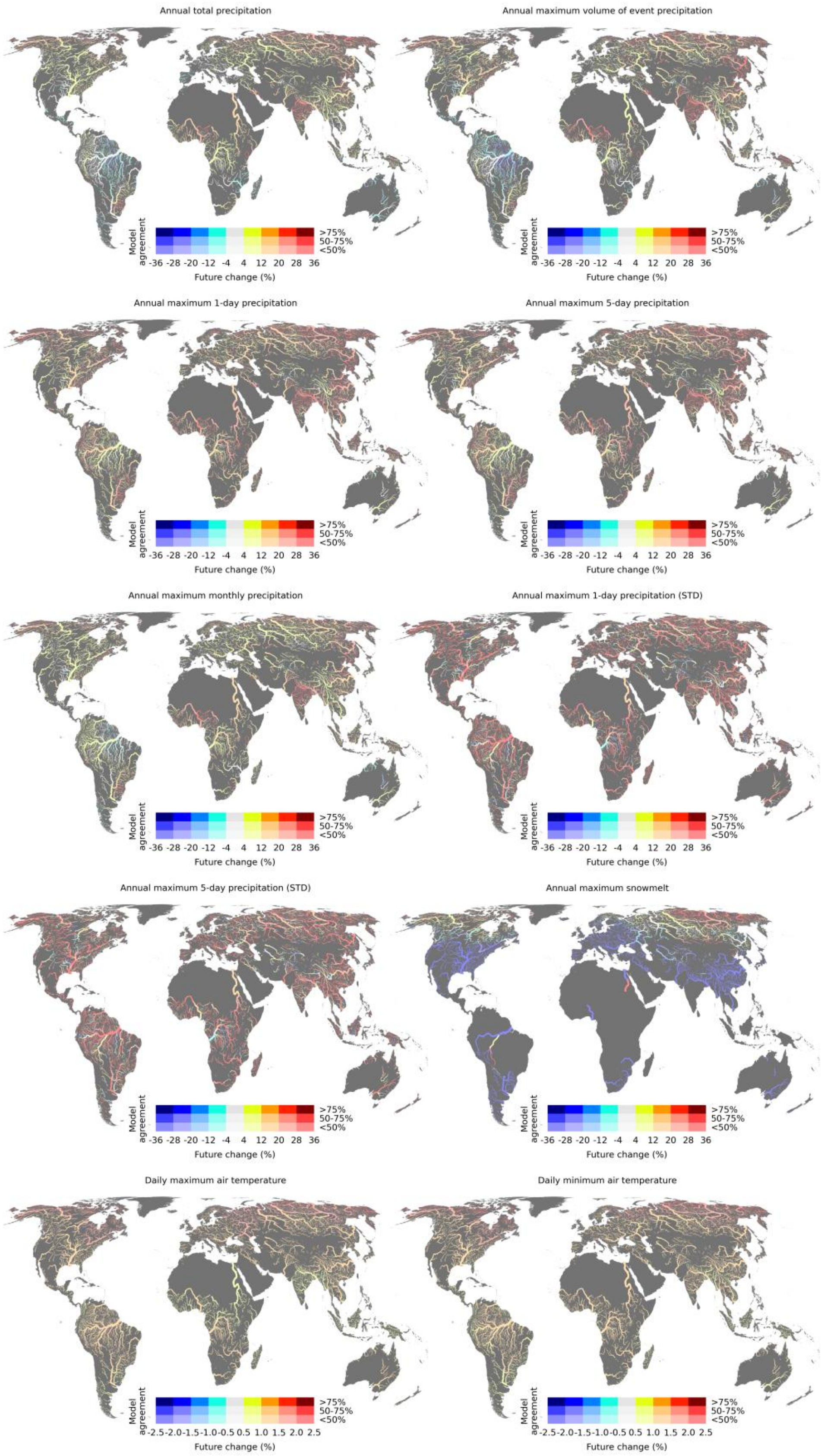


Fig. S11 Multi-model ensemble mean in the projected changes of each climate predictor under SSP2-4.5. The colour hues in the maps show the magnitude of change and the saturation indicates the agreement, among ensemble members, in the sign of change.

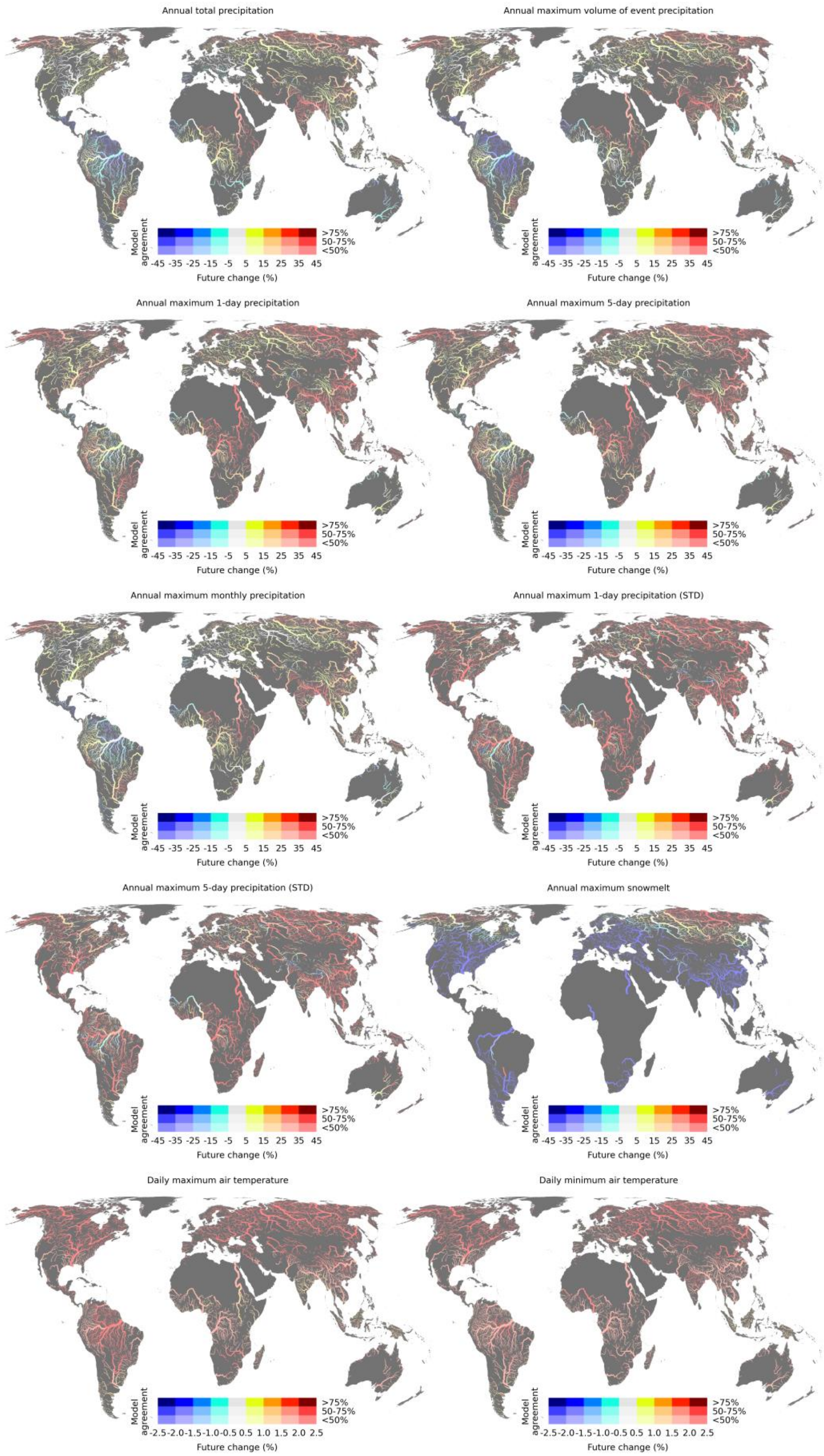


Fig. S12 Same as Fig. S11 but for the difference between SSP2-4.5 and SSP5-8.5 scenarios.

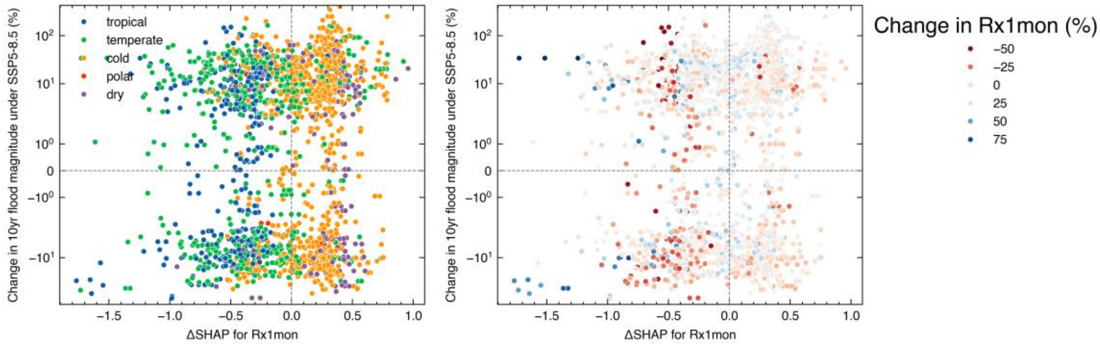


Fig. S13 We compare a simulation-only ML model (*$ML_{cmip6\text{-}only}$*) trained on climate model simulations and a perfect prediction ML model (*$ML_{perfect}$*) in which the most important CMIP6 climate predictor is replaced by a "perfect" predictor calculated from the MSWEP v2 dataset. The difference in SHAP values (i.e., ΔSHAP) of the annual maximum 1-month precipitation (Rx1mon) generated from the *$ML_{cmip6\text{-}only}$* and *$ML_{perfect}$* models are compared with the projected changes of flood magnitude for the 7,115 river basins under SSP5-8.5. Multi-model ensemble means in the projected changes of flood magnitude and Rx1mon are calculated between historical (1981−2014) and future (2071−2100) periods. The dots in the right panel are coloured by the projected changes of Rx1mon.

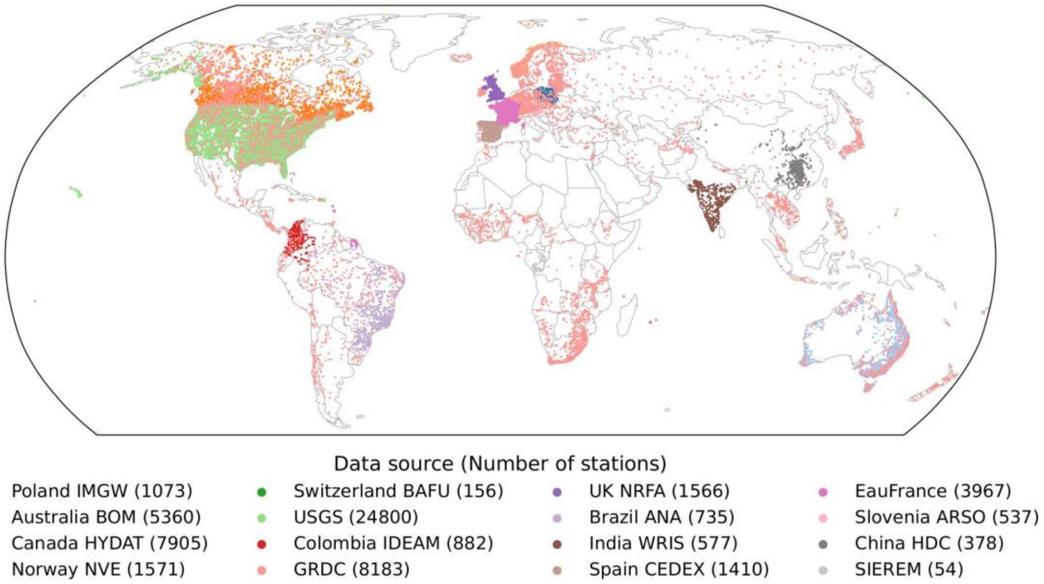
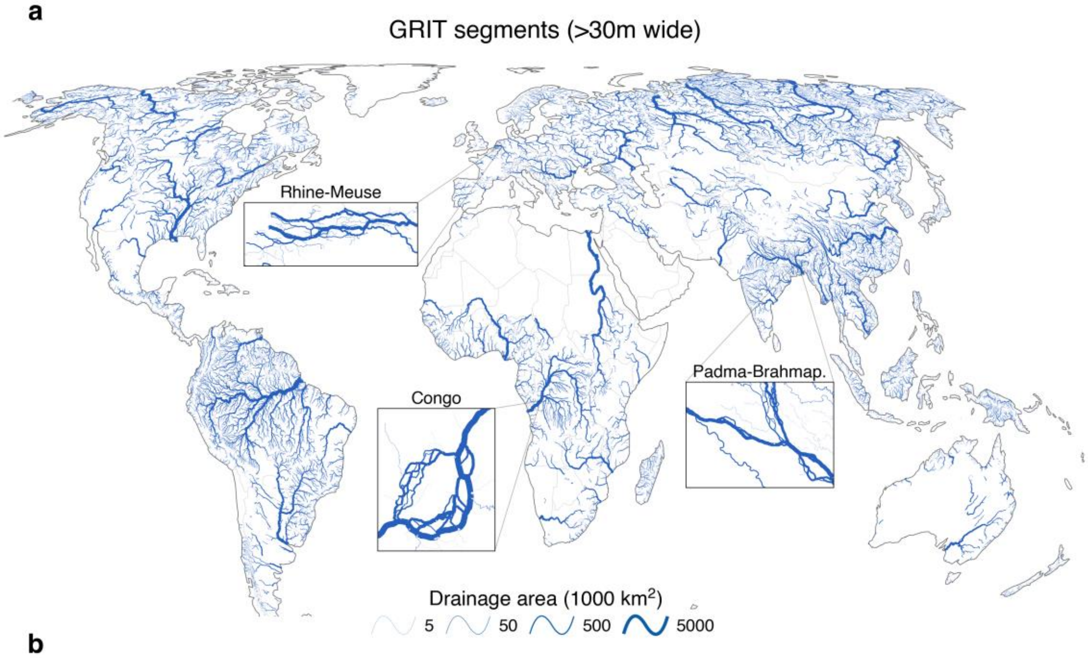


Fig. S14 Spatial distribution of global streamflow stations and data sources.

a



b

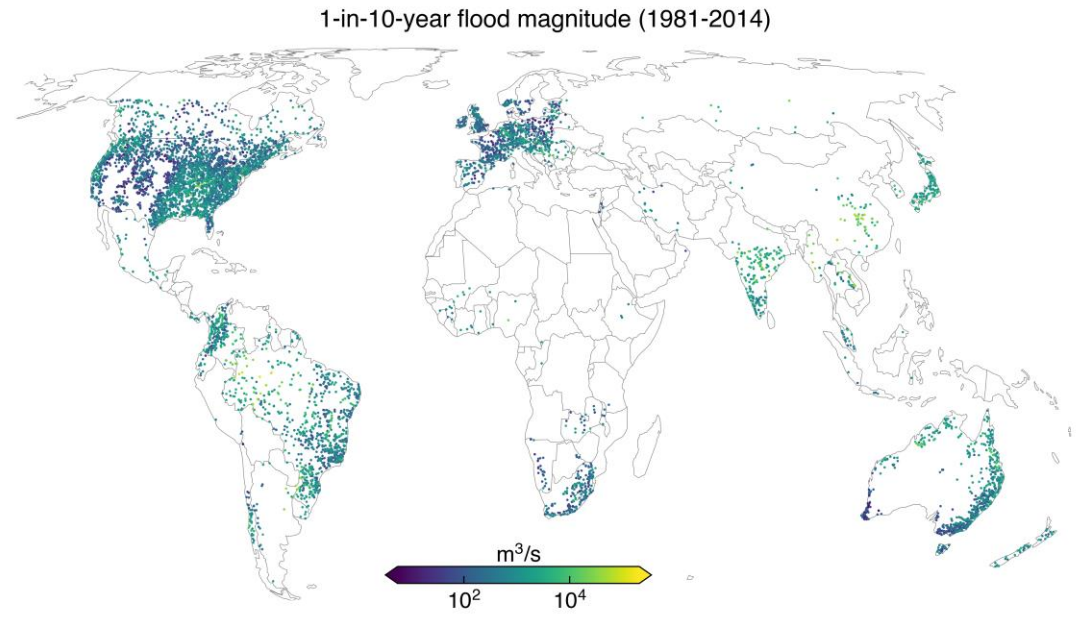


Fig. S15 GRIT reaches and 10-year flood magnitude. **a** presents reaches from the GRIT network with Global River Widths from Landsat (GRWL) width greater than 30 m as well as their downstream segments. The width is scaled by partitioned drainage area here for visualization purposes. The inset rectangles show the detail of the multi-threaded river network in three different regions of the globe (Rhine-Meuse, Congo, Padma-Brahmaputra). **b** presents the 10-year flood magnitude estimated using the empirical probability function of annual maximum discharge for each gauge, where each gauge has at least 10 years of daily discharge observations during the period 1981−2014.

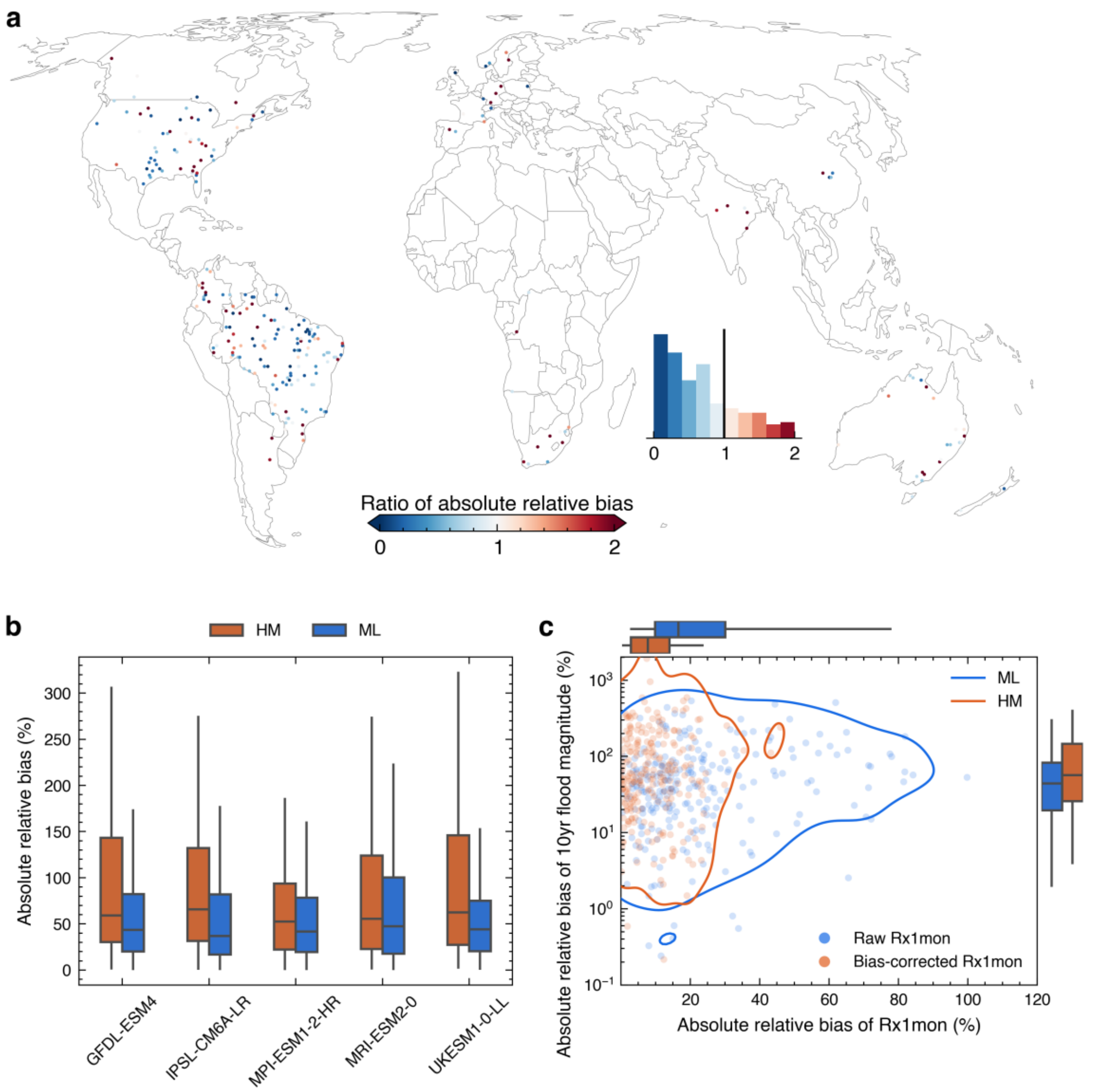


Fig. S16 Relative bias of the ML models versus hydrological models from the ISIMIP3b experiment in predicting the 10-year flood event during the period 2001−2014. Here grid cells are limited to those where the drainage area calculated using DDM30 differed by no more than 20% from the drainage area calculated using river basin polygons from the GRIT dataset. **a** Ratio of absolute relative bias based on multi-model ensemble mean predictions from the 19-GCM ML models and 10 hydrological models (HM) from the ISIMIP3b experiment; inset histogram indicates the distribution of absolute relative biases. **b** Absolute relative biases generated from the ML and HM models driven by the same 5 GCMs (using the 5 GCMs from the ISIMIP3b experiment). **c** CMIP6 bias of the annual maximum 1-month precipitation (Rx1mon) and the bias of 10-year flood magnitude generated by the ensemble mean 19 ML (blue) and 10 HM (red) predictions; The contour lines in **c** represent the 95% confidence contour, indicating that 95% of the probability mass lies within the contour. The whiskers of boxplots extend to the 2.5% and 97.5% percentiles. Boxplots at the right and top of panel **c** indicate the distribution of flood biases and Rx1mon biases, respectively. Note: average 10-year flood magnitude is computed as specific discharge (mm/day) for the HMs in each grid cell. The ML models are calibrated during the period 1981−2000, while the HM models are calibrated during the period 1961−1970[51].

References

1. Kang, S. *et al.* Observation-Constrained Projection of Flood Risks and Socioeconomic Exposure in China. *Earth's Futur.* **11**, (2023).
2. Di Sante, F., Coppola, E. & Giorgi, F. Projections of river floods in Europe using EURO-CORDEX, CMIP5 and CMIP6 simulations. *Int. J. Climatol.* **41**, 3203–3221 (2021).
3. Hirabayashi, Y., Tanoue, M., Sasaki, O., Zhou, X. & Yamazaki, D. Global exposure to flooding from the new CMIP6 climate model projections. *Sci. Rep.* **11**, (2021).
4. Jiang, R. *et al.* Substantial increase in future fluvial flood risk projected in China's major urban agglomerations. *Commun. Earth Environ.* **4**, (2023).
5. Liu, W. *et al.* Observation-constrained projection of global flood magnitudes with anthropogenic warming. *Water Resour. Res.* **57**, (2021).
6. Xuan Do, H. *et al.* Historical and future changes in global flood magnitude - evidence from a model-observation investigation. *Hydrol. Earth Syst. Sci.* **24**, 1543–1564 (2020).
7. Yamazaki, D., Watanabe, S. & Hirabayashi, Y. Global Flood Risk Modeling and Projections of Climate Change Impacts. *Geophys. Monogr. Ser.* **233**, 185–203 (2018).
8. Wang, H. *et al.* Anthropogenic climate change has influenced global river flow seasonality. *Science (80-.).* **383**, 1009–1014 (2024).
9. Alfieri, L. *et al.* Global projections of river flood risk in a warmer world. *Earth's Futur.* **5**, 171–182 (2017).
10. J.-S., N. & Mishra, V. Projected increase in widespread riverine floods in India under a warming climate. *J. Hydrol.* **630**, (2024).
11. Wobus, C. *et al.* Climate change impacts on flood risk and asset damages within mapped 100-year floodplains of the contiguous United States. *Nat. Hazards Earth Syst. Sci.* **17**, 2199–2211 (2017).
12. Sharma, S., Gomez, M., Keller, K., Nicholas, R. E. & Mejia, A. Regional flood risk projections under climate change. *J. Hydrometeorol.* **22**, 2259–2274 (2021).
13. Bates, P. D. *et al.* Combined modeling of US fluvial, pluvial, and coastal flood hazard under current and future climates. *Water Resour. Res.* **57**, (2021).
14. Archer, L. *et al.* Current and future rainfall-driven flood risk from hurricanes in Puerto Rico under 1.5 and 2 °C climate change. *Nat. Hazards Earth Syst. Sci.* **24**, 375–396 (2024).
15. Archer, L. *et al.* Future change in urban flooding using new convection-permitting climate projections. *Water Resour. Res.* **60**, (2024).
16. Kim, H. & Villarini, G. Higher emissions scenarios lead to more extreme flooding in the United States. *Nat. Commun.* **15**, (2024).
17. Wasko, C., Guo, D., Ho, M., Nathan, R. & Vogel, E. Diverging projections for flood and rainfall frequency curves. *J. Hydrol.* **620**, (2023).
18. Zhang, S. *et al.* Reconciling disagreement on global river flood changes in a warming climate. *Nat. Clim. Chang.* **12**, 1160–1167 (2022).
19. Swarnkar, S. & Mujumdar, P. Increasing flood frequencies under warming in the west-central Himalayas. *Water Resour. Res.* **59**, (2023).
20. Bi, D. *et al.* The ACCESS coupled model: Description, control climate and evaluation. *Aust. Meteorol. Oceanogr. J.* **63**, 41–64 (2013).
21. Law, R. M. *et al.* The carbon cycle in the Australian Community Climate and Earth System Simulator (ACCESS-ESM1) - Part 1: Model description and pre-industrial simulation. *Geosci. Model Dev.* **10**, 2567–2590 (2017).

22. Swart, N. C. *et al.* The Canadian Earth System Model version 5 (CanESM5.0.3). *Geosci. Model Dev.* **12**, 4823–4873 (2019).
23. Séférian, R. *et al.* Evaluation of CNRM Earth System Model, CNRM-ESM2-1: Role of Earth System Processes in Present-Day and Future Climate. *J. Adv. Model. Earth Syst.* **11**, 4182–4227 (2019).
24. Voldoire, A. *et al.* Evaluation of CMIP6 DECK Experiments With CNRM-CM6-1. *J. Adv. Model. Earth Syst.* **11**, 2177–2213 (2019).
25. Shi, X., Lohmann, G., Sidorenko, D. & Yang, H. Early-Holocene simulations using different forcings and resolutions in AWI-ESM. *Holocene* **30**, 996–1015 (2020).
26. Massonnet, F. *et al.* Replicability of the EC-Earth3 Earth system model under a change in computing environment. *Geosci. Model Dev.* **13**, 1165–1178 (2020).
27. Döscher, R. *et al.* The EC-Earth3 Earth system model for the Coupled Model Intercomparison Project 6. *Geosci. Model Dev.* **15**, 2973–3020 (2022).
28. Dunne, J. P. *et al.* The GFDL Earth System Model version 4.1 (GFDL-ESM 4.1): Overall coupled model description and simulation characteristics. *J. Adv. Model. Earth Syst.* **12**, (2020).
29. Volodin, E. & Gritsun, A. Simulation of observed climate changes in 1850-2014 with climate model INM-CM5. *Earth Syst. Dyn.* **9**, 1235–1242 (2018).
30. Andrews, M. B. *et al.* Historical Simulations With HadGEM3-GC3.1 for CMIP6. *J. Adv. Model. Earth Syst.* **12**, (2020).
31. Tatebe, H. *et al.* Description and basic evaluation of simulated mean state, internal variability, and climate sensitivity in MIROC6. *Geosci. Model Dev.* **12**, 2727–2765 (2019).
32. Hajima, T. *et al.* Development of the MIROC-ES2L Earth system model and the evaluation of biogeochemical processes and feedbacks. *Geosci. Model Dev.* **13**, 2197–2244 (2020).
33. Müller, W. A. *et al.* A Higher-resolution Version of the Max Planck Institute Earth System Model (MPI-ESM1.2-HR). *J. Adv. Model. Earth Syst.* **10**, 1383–1413 (2018).
34. Mauritsen, T. *et al.* Developments in the MPI-M earth system model version 1.2 (MPI-ESM1.2) and its response to increasing CO2. *J. Adv. Model. Earth Syst.* **11**, 998–1038 (2019).
35. Yukimoto, S. *et al.* The meteorological research institute Earth system model version 2.0, MRI-ESM2.0: Description and basic evaluation of the physical component. *J. Meteorol. Soc. Japan* **97**, 931–965 (2019).
36. Seland, Ø. *et al.* Overview of the Norwegian Earth System Model (NorESM2) and key climate response of CMIP6 DECK, historical, and scenario simulations. *Geosci. Model Dev.* **13**, 6165–6200 (2020).
37. Sellar, A. A. *et al.* UKESM1: Description and Evaluation of the U.K. Earth System Model. *J. Adv. Model. Earth Syst.* **11**, 4513–4558 (2019).
38. Boucher, O. *et al.* Presentation and Evaluation of the IPSL-CM6A-LR Climate Model. *J. Adv. Model. Earth Syst.* **12**, (2020).
39. ESA. Land Cover CCI Product User Guide Version 2.0. http://maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-Ph2-PUGv2_2.0.pdf (2017).
40. Lehner, B., Verdin, K. & Jarvis, A. New global hydrography derived from spaceborne elevation data. *Eos (Washington. DC).* **89**, 93–94 (2008).
41. Pelletier, J. D. *et al.* A gridded global data set of soil, intact regolith, and sedimentary deposit thicknesses for regional and global land surface modeling.

J. Adv. Model. Earth Syst. **8**, 41–65 (2016).
42. Huscroft, J., Gleeson, T., Hartmann, J. & Börker, J. Compiling and mapping global permeability of the unconsolidated and consolidated earth: GLobal HYdrogeology MaPS 2.0 (GLHYMPS 2.0). *Geophys. Res. Lett.* **45**, 1897–1904 (2018).
43. Zomer, R. J., Xu, J. & Trabucco, A. Version 3 of the Global Aridity Index and Potential Evapotranspiration Database. *Sci. Data* **9**, (2022).
44. Beck, H. E. *et al.* Present and future köppen-geiger climate classification maps at 1-km resolution. *Sci. Data* **5**, (2018).
45. Beck, H. E. *et al.* MSWep v2 Global 3-hourly 0.1° precipitation: Methodology and quantitative assessment. *Bull. Am. Meteorol. Soc.* **100**, 473–500 (2019).
46. Harrigan, S. *et al.* GloFAS-ERA5 operational global river discharge reanalysis 1979-present. *Earth Syst. Sci. Data* **12**, 2043–2060 (2020).
47. Baker, D. B., Richards, R. P., Loftus, T. T. & Kramer, J. W. A new flashiness index: Characteristics and applications to Midwestern rivers and streams. *J. Am. Water Resour. Assoc.* **40**, 503–522 (2004).
48. Tatem, A. J. WorldPop, open data for spatial demography. *Scientific Data* vol. 4 at https://doi.org/10.1038/sdata.2017.4 (2017).
49. Kummu, M., Taka, M. & Guillaume, J. H. A. Gridded global datasets for Gross Domestic Product and Human Development Index over 1990-2015. *Sci. Data* **5**, (2018).
50. 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**, (2023).
51. Yoshida, T. *et al.* Inference of parameters for a global hydrological model: Identifiability and predictive uncertainties of climate-based parameters. *Water Resour. Res.* **58**, (2022).